National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

NL-339 (r. 88/04) c

Canada

The University of Alberta

In Search of Breast Carcinoma: A Mammographic Asymmetry Approach

by

Tin-Kit Lau

A thesis
submitted to the Faculty of Graduate Studies and Research
in partial fulfillment of the requirements for the degree
of Master of Science

Department of Computing Science

Edmonton, Alberta
Spring 1990

National Library of Canada
Bibliothèque nationale du Canada

Canadian Theses Service
Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

ISBN 0-315-60282-1

NL-339 (r. 88/04) c

Canada

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: Tin-Kit Lau

TITLE OF THESIS: In Search of Breast Carcinoma:
A Mammographic Asymmetry Approach

DEGREE: Master of Science

YEAR THIS DEGREE GRANTED: 1990

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

(Signed) .....

Permanent Address:
Flat 8, 6/Fl., Block C,
Hilton Plaza,
7 Sha Tin Center Street,
Sha Tin, N.T.,
Hong Kong

Date: March 9, 1990

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled **In Search of Breast Carcinoma: A Mammographic Asymmetry Approach** submitted by **Tin-Kit Lau** in partial fulfillment of the requirements for the degree of Master of Science.

Dr. Walter F. Bischof (Co-supervisor)

Dr. Xiaobo Li (Co-supervisor)

Dr. Wayne Davis

Dr. Jia You

Dr. Roger Toogood (External)

Date: Jan. 17, 1990

To: My Parents
and
my family ...

Abstract

This thesis describes a system for automated detection of breast cancers in mammograms. It uses the "asymmetry approach", which suspects the presence of a breast tumor at all locations, where there is a significant difference in appearance between corresponding locations in the left and right breast. The difficulty of this approach lies in the fact that these differences have to be found against a background of naturally occurring asymmetries of healthy breast tissue. A complete detection method is implemented and presented in this thesis, with each step discussed in detail.

After the film mammograms have been digitized, a global thresholding segmentation technique is used to extract the breast areas to be used for comparison. A cubic B-spline corner detection method is used to detect control points, namely the nipple and intersection points between breast boundary and back boundary. These control points are then incorporated in a geometric transformation which aligns the pair of mammogram images.

Four different measurements are studied for asymmetry measurement. They include locally normalized digital subtraction, locally normalized variance difference, brightness-to-roughness ratio difference and directionality measures. Each of these measurements is described individually and the formation of an overall mammographic asymmetry measurement from these four measurements, using a vertical directionality measure as weighting factor, is discussed.

A two-stage thresholding method is used to select suspicious areas from the response image produced by the mammographic asymmetry measurement. For a set of ten test cases, this method correctly identified twelve out of the thirteen suspicious tumor locations marked by radiologists.

Acknowledgements

First of all, I wish to express my sincere gratitude to my supervisors, Dr. Walter F. Bischof and Dr. Xiaobo Li, for their guidance, support and encouragement throughout the research of this thesis. I would like to specially thank Dr. Bischof for his invaluable suggestions and timely discussion on the problems encountered in the course of this research.

I would also like to thank the members of the examining committee: Dr. Wayne Davis, Dr. Jia You, and Dr. Roger Toogood for their time in reviewing this thesis and their helpful comments. Special thanks to Dr. Bill Castor, director of radiology department of the Cross Cancer Institute, for providing mammograms for this project and sharing his vast knowledge in mammogram interpretation with me.

I am also grateful to Mr. Vickitt Lau for his inspiration during the initial period of this degree program. I would also wish to thank Mr. and Mrs. Hing for making my stay in Edmonton possible and a pleasant one. I am also indebted to Mr. Kris Ng for his help in compiling this thesis. I would like to thank Mrs. Sharon Darke for her time in proofreading this thesis. Thanks are also due to Mr. Anthony Ng for his support and time in discussion throughout the course of this thesis's research.

Most of all, I wish to express my utmost gratitude to my family and Miss Ping Yeung for their love, patience, encouragement and continuous support. The completion of this thesis would not have been possible without their constant support and encouragement.

This research was financially supported in part by the Department of Computing Science, University of Alberta, and the National Science and Engineering Research Council of Canada.

Table Of Contents

1 Introduction 1
- 1.1 Overview on Breast Cancer . . . 2
- 1.2 Overview on Breast Imaging Techniques . . . 3
- 1.3 Mammographic Diagnosis . . . 7
  - 1.3.1 Mammogram Quality . . . 9
  - 1.3.2 Interpretation Procedure . . . 10
- 1.4 Problem Statement . . . 11
- 1.5 Conventions and Notations . . . 12
- 1.6 Thesis Organization . . . 14

2 Review on Automated Breast Tumor Detection 16
- 2.1 Difficulties in Detecting Breast Tumor from Mammograms . . . 20
- 2.2 Advantages of Asymmetry Approach . . . 22
- 2.3 Processes of Asymmetry Approach . . . 23
- 2.4 Examples of Mammogram Images . . . 24

3 Preprocessing of Mammograms 27
- 3.1 Digitization of Mammograms . . . 28
- 3.2 Image Enhancement . . . 30
- 3.3 Area of Interest Detection . . . 30
- 3.4 Control Point Extraction . . . 35
  - 3.4.1 Difficulty in Automating Extraction . . . 36
  - 3.4.2 Proposed Approach . . . 36
  - 3.4.3 Breast Boundary Extraction and Smoothing . . . 38
  - 3.4.4 Cubic B-Spline Method . . . 39

3.5 Geometric Transformation 43
3.5.1 Difficulties In Obtaining Accurate Method 43
3.5.2 Review on Existing Methods 44
3.5.3 Proposed Method 46

4 Detection of Mammographic Asymmetry **56**
4.1 Digital Subtraction 57
4.2 Analysis Using Texture Statistics 59
4.3 Roughness (Variance) 64
4.3.1 Globally Normalized Variance 65
4.3.2 Locally Normalized Variance 67
4.4 Brightness-to-Roughness Ratio 70
4.5 Directionality 72
4.5.1 Fourier Transform 73
4.6 Formation of Asymmetry Measurement 77
4.7 Selection of Suspicious Areas 79
4.7.1 Experimental Results 81

5 Conclusions and Future Research **90**
5.1 Future Directions 94

Bibliography **96**

Glossary **101**

List of Figures

1.1 Coordinate System for Digital Mammogram Image 13

2.1 Mammographic Asymmetry Approach Processes 25

3.1 Boundary diagram of digitized mammogram 29
3.2 Histogram for Plate 2.1 33
3.3 Control points of a typical breast area 37
3.4 Cubic B-spline approximation for P_i to P_{i+1} 40
3.5 Cubic B-spline approximation for P'_i to P'_{i+1} 41
3.6 Shearing effect of breast compression 45
3.7 Subdivision of breast area 48
3.8 Elimination of skewing effect 49
3.9 Transformation of sub-region L 50
3.10 Bilinear transformation of non-reference image 51

4.1 Window subdivision for locally normalized variance. 68
4.2 Subdivision of power spectrum for directionality. 75

List of Tables

4.1 Comparison of Brightness and Roughness on various breast tissue. 70
4.2 Comparison of performance of different q_2 value in second stage thresholding procedure. 84
4.3 Comparison of computer's and radiologist's interpretation on ten test cases. 84

List of Plates

2.1 Medio-Lateral View 26
2.2 Cranio-Caudal View 26
3.1 Binary image for the left image of Plate 2.1 53
3.2 Thresholded averaging of binary image in Plate 3.1 53
3.3 Boundary detected for the left image of Plate 2.1 54
3.4 Control points detected for Plate 3.3 54
3.5 Geometrically transformed images for Plate 2.1 55
4.1 Result of Locally Normalized Digital Subtraction 85
4.2 Result(case:49746cc) of Locally Normalized Digital Subtraction 85
4.3 Result of Globally Normalized Variance Difference 86
4.4 Result of Locally Normalized Variance Difference 86
4.5 Result of Brightness-to-Roughness Ratio Difference 87
4.6 Result of Directionality Difference 87
4.7 Result of Combined Responses on test case: Plate 2.1 88
4.8 Final Result of Suspicious Areas Detected on Case:19010ML . 88
4.9 Final Result of Suspicious Areas Detected on Case:19010CC . 89

Chapter 1

Introduction

The rising incidence of breast cancer shows that a mass mammography screening program is in great demand. Mammography is a technically demanding procedure that requires experienced, skillful, and meticulous interpretation [McL84]. A mass screening program requires interpretation of a large number of images by expertly trained radiologists. The cost and shortage of mammography experts have been the major deterrents of its widespread acceptance and implementation. With advanced computer technology in biomedical image processing, it is possible and feasible to automate some prescreening processes. This would not only reduce the cost-benefit ratio but will also enable radiologists to concentrate on suspicious cases. This thesis attempts to search for algorithms to automate the breast cancer detection process using the mammographic asymmetry approach.

1.1 Overview on Breast Cancer

Breast cancer has become a predominant fear among women today. The chance of developing this disease is so high that about nine percent of the female population (1 in 11 born) in United States will develop breast cancer in their lifetime. More than 25% of all cancers diagnosed in women are breast cancers and they account for 19% of all cancer deaths [SK83].

During the past decade, there has been a determined effort to identify the factors capable of separating women at high risk from those with little or no likelihood of developing this disease. For example, in younger women, factors might lean more to familial factors, while those in older women to environmental conditions. Other indicators such as personal/family history of breast cancer, hormonal status, substantial doses of ionizing radiation, body weight and mammographic pattern are also suggested; but the etiology of the disease is still unknown [Les84]. Despite the fact that primary prevention is still not possible for the above reasons, detection of breast cancer in the early stage can prevent it from becoming incurable and reduce mortality. Thus, early detection is the best hope for improved survival [McL84].

It is well appreciated that most breast cancers cannot be palpated, i.e. detected externally by physical examination, until they have achieved a size of about $1\,cm^3$, or 1 billion (10^9) cells. Many kinetic studies have indicated that this requires about 30 population doubles. When it is recognized that a

doubling time might take 30–200 or more days, it becomes apparent that a tumor that is regarded as clinically early, is in fact biologically late, requiring only 10–20 more doubles before causing the death of the host [Fis85]. According to previous studies, about 80% of *Stage I* and 95% of *Stage 0* patients achieve 10 years disease-free survival rate after either conservation surgery and radiation or radical/modified-radical mastectomy alone [Les84]. *Stage I* patients are those with a tumor of 2 cm or less in greatest diameter, without axillary nodal or distant metastases involvement according to the America Joint Committee on Cancer (AJCC)'s TNM breast cancer staging system. *Stage 0* is not an official part of the AJCC's system; but is commonly used by some authors in referring minimal, *in situ* cancer of 5 mm or less in size. Thus, it is important for medical experts to be able to detect breast cancer in *Stage 0*. Mammography is the only method to date which has convincingly demonstrated the ability to detect breast cancer at an early stage with high sensitivity and specificity [TD87].

1.2 Overview on Breast Imaging Techniques

To detect early stage breast cancer in mass screening process, non-invasive detection methods must be developed and used when a physical examination method is no longer adequate to detect cancer at the nonpalpable stage. Many breast imaging techniques have been developed for this purpose. Each

technique has its own merit as well as its own drawbacks. In the following few paragraphs, a few techniques will be discussed briefly.

Ultrasound is a breast imaging technique where pulsed mechanical waves of 2.25–10 megahertz are directed through the patient's breast which is immersed in a coupling agent of water, mineral oil, or acoustic gel. When the pulsed wave encounters an interface with a difference in acoustic impedance, such as the wall of a cystic structure, a small portion of the beam is reflected back to the transducer. These reflected echoes are then displayed on a scan convertor for viewing or recording [Col87]. Ultrasound has not only the ability to image the breast repeatedly with no known deleterious effects and differentiate solid mass from cystic mass, but also finding masses in dense breasts that are difficult to diagnose by X-ray mammography. However, the equipment is very expensive, examination time is relatively long and resolution is limited (lesions smaller than 1 cm remain usually undetected) [Mar83].

Computed Tomography (CT Scanning) is another breast imaging technique in which a scanner of finely collimated x-ray beams opposite either a single or multiple detectors arranged radially in a 360° circle is used. A typical 8 mm cross-sectional slice of the patient's breast is scanned by pulsing the x-ray beam over a 360° sweep. The cross-sectional image based on the various absorption coefficients measured by the detectors is then reconstructed and displayed on a screen [GKR77, DS83] . This method shows a high degree

of accuracy in detecting breast cancers and is comparable, if not superior, to X-ray mammography. However, adverse reactions from the intravenous contrast medium, high costs and lengthy procedure make it inappropriate as a screening tool for the general population [CSF+80].

Thermography, a method for producing an infrared image of the surface of an anatomic section, is based on the principle that the amount of radiation emitted by an object depends on its absolute temperature. Infrared sensitive detectors produce an image representing the surface temperature of the skin [DMH76]. This method has also been used for breast cancer detection as the temperature of the skin in the vicinity of mammary cancers is generally higher than the corresponding area of the opposite breast. The temperature-sensing method has no radiation hazard and is inexpensive to perform; however, it does not give consistent results for various reasons, such as early stage tumors (less than 2 cm) do not show sufficient temperature discrepancy from normal breast tissue and also due to the large variation in metabolic activities in different types of tumors [LR80, Mar83].

In addition to the above mentioned techniques, there are also transillumination, sonography, magnetic resonance imaging, etc.; but these methods have neither proven their effectiveness in extensive clinical studies nor have they been proven adequate for mass screening.

X-ray mammography has not only been used as a breast tumor screening method for more than two decades, but has also proven its effectiveness

in detecting early stage tumors in various mass screening studies, such as BCDDP[1] and HIP[2]. It has been shown that this method successfully detected 96% of breast cancer cases and that 48% of those detected cases are at a nonpalpable stage [ST82]. But with no exception, this technique also has its own drawback, such as radiation hazards and insufficient contrast to show individual structures within breast tissue. During the past 20 years, new improved equipment and techniques, such as breast compression and film-screen combination, have enabled good quality mammograms to be produced while radiation dosage is minimized. Experience gained also improves the accuracy in interpreting mammograms by radiologists. All these improvements have made mammography the standard for diagnosis of breast cancer as no other method approaches its effectiveness [Mar83].

Based on the consideration of cost-beneficial ratio, effectiveness, availability and length of examination; mammography is ideal, if not the best, screening method for detecting breast tumor. In fact, it is currently the most common screening procedure used across the nation.

---

[1]Breast Cancer Detection Demonstration Project supported by the National Institute of Health an American Cancer Society in 1974–1981

[2]Controlled screening trial for breast cancer conducted by Health Insurance Plan of New York in 1963–1969

1.3 Mammographic Diagnosis

At the present time, X-ray mammography is well accepted as the most reliable and effective method for diagnosing early breast tumors. Combined with meticulous interpretation by expert radiologists, X-ray mammography demonstrates a high degree of accuracy in detecting early stage breast tumors. To automate the diagnosis procedure of mammograms, it is important to have fundamental knowledge in mammography techniques and its diagnostic procedures.

Two types of mammographic systems commonly used today are *film mammography* and *xeromammography*. The main difference between these two lies in the image recording process. In film-mammography, conventional fine grain direct exposure emulsion film is used to produce mammograms with exceptional radiographic details. The screen-film technique which combines a calcium tungstate intensifying screen in intimate contact with a single emulsion film by a vacuum further reduces radiation dose applied to the patient. In addition, this technique yields mammographic images that are far superior to non-screen film mammograms [LJ87].

In xeroradiography, an aluminum plate coated with photoconductor, selenium, is the counterpart of the conventional X-ray film. A latent image in the form of a pattern of electrical charges is produced on a uniformly charged selenium plate after exposure to X-rays. This pattern of electrical charges on

the plate is then transcribed into a visible image on a plastic-coated paper [Fei87]. Unlike film-mammography which produces photographic negative images, xeromammography produces positive images with the advantage of eliminating the need for a view box in the interpretation process.

Both systems have their own advantages over each other. Xeromammogram is characterized by its wide recording latitude which yields good contrast and detail throughout a wide range of tissue density and thickness [Fei87]. Also, the xeroradiographic edge enhancement property accentuates the visualization of margins of breast masses, spiculations and calcific particles [Fei87]. On the other hand, screen-film mammograms have a higher overall difference in density between a mass and surrounding tissue (*board area contrast*) than xeromammograms. Radiation dose (26–28 kVp) applied is also lower when compared to xeromammography (45–55 kVp) [Fei87].

Despite all the differences between the above two mammographic systems, studies do not show significant difference in their ability in detecting breast cancer. Instead, the success of either mammographic system depends on the skill and experience of the radiologist who operates them [GB83].

In the research of this thesis, screen-film mammograms are provided by the Cross Cancer Institute of Alberta and are used in the detection program.

1.3.1 Mammogram Quality

Mammograms are the primary source of data of the entire diagnostic procedure. High quality mammograms are required for ensuring a high detection rate. The quality of mammogram basically depends on two major factors: mammographic equipment and mammographer's skill and experience.

Mammographic equipment has been improved dramatically over the past 20 years. Dedicated mammographic machines are built to produce good quality mammograms while minimizing radiation doses. Various compression equipment have been developed for proper breast compression to a more uniform thickness; thus allowing a more even X-ray penetration and reducing the amount of scattered radiation. Moreover, the highly automated exposure control and the uniformly high quality screen-film combinations make latent images of similarly high quality of other radiographic films [TD85].

Radiologists' skill and experience acquired through training and clinical practice have been proven invaluable in assisting patients for proper positioning of breast to be imaged. They also ensure that the breast is properly compressed and immobilized during the course of mammogram being taken; hence, eliminating any blurring effect that might be caused by movement. Afterall, it is the radiologist's responsibility to ensure that all components of the mammography system are functioning correctly to achieve a high-quality mammogram [TD85].

1.3.2 Interpretation Procedure

Mammographic examination routinely includes two views of each breast, namely *medio-lateral* view and *cranio-caudal* view. Auxiliary views are sometimes required; but very few cancers are located in the breast in such a way as to require extra and extended views beyond standard medio-lateral and cranio-caudal views [ST82]. Hence, we will focus on the two standard views only.

The film reading procedure performed by expert radiologists involves the interpretation of each individual film and a comparison of identical views from both breasts by using various masking techniques. Breast carcinomas can be pathologically classified into four major categories: *Circumscribed masses*, *Stellate lesions*, *Micro-calcifications* and *Thickened skin syndrome*. Each type of tumor presents a certain number of discriminating mammographic signs which can be used as detection criteria on individual film. For example, (i) circumscribed masses have hazy and indistinct boundaries, approximately circular in shape and may have short spicules on their surfaces; (ii) stellate lesions have distinct central mass surrounded by a radiating corona of spicules. In addition to these primary signs, there are also mammographic signs to differentiate benign and malignant tumors. For instance, boundary sharpness and density are used in circumscribed masses and consistency in appearance across projections for stellate lesions [TD85]. Furthermore, other mammo-

graphic signs such as retraction, localized skin thickening are also used in detecting tumors.

Apart from the method of detecting breast tumors using the above mentioned criteria, comparison of identical view mammograms of both breasts to find asymmetries is also an effective method, not only for reducing the search area but also for the detection of breast cancers. Comparison of corresponding regions of two breasts is facilitated by using the masking technique; i.e. covering specific regions of the breast and viewing the uncovered portions [TD85]. Human breasts are approximately symmetric and there is a very low likelihood that a patient has the *SAME* (size, type) tumor growth at exactly the same location in both breasts. The earliest sign of tumor may be distortion of the breast structure which will produce asymmetries between both breast images. When these architectural disturbances are recognized, hidden lesions can be detected [TD85].

1.4 Problem Statement

This thesis investigates methods to detect early stage breast carcinomas in the mammography screening process, using the mammographic asymmetry approach. This approach is based on the fact that human breasts are approximately symmetric and there is a low likelihood that tumors of the same type and size occur at exactly the same location in both breasts at the same time.

The existence of a tumor in one breast produces an asymmetry in terms of brightness, contrast and distortion of breast structure at the tumor location on identical views of both breasts. Detection of any asymmetry on identical views of both breasts can lead to the detection of hidden lesion.

The problem can be stated as: *Given a set of identical view mammograms (either medio-lateral or cranio-caudal), the goal is to develop a detection method based on the mammographic asymmetry approach which determines the location and size of all suspicious areas.* It must be noted that suspicious areas include not only tumor areas but also areas that expert radiologists would choose for further detailed examination. No attempt is made at classifying detected suspicious areas into tumor categories as this is beyond the scope of this thesis.

1.5 Conventions and Notations

Digital format of mammogram images are used in the detection program of this thesis. Mammogram films are digitized into 512x512 image elements or pixels by using of a TV camera. The grey level (brightness level) of each image element or pixel is represented by an integer and is stored in a two dimensional array in which the corresponding row and column indices identify the position of the pixel within the image.

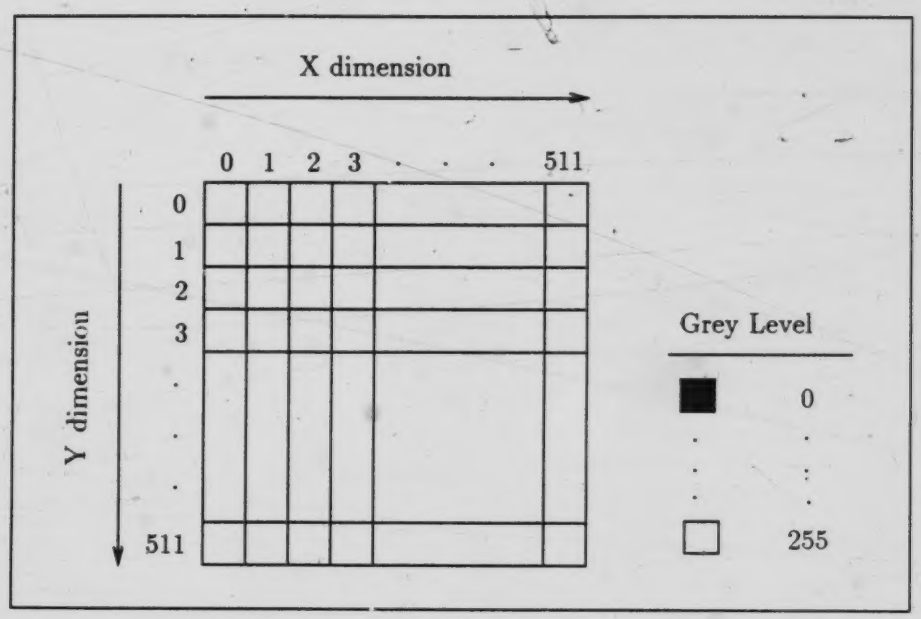


Figure 1.1: Coordinate System for Digital Mammogram Image

A total of 256 grey levels are used and numbered from 0 to 255, with 0 representing black and 255 representing white. The coordinate system has the origin (0,0) located at the upper-left hand corner of the image and each pixel is identified by its position (x,y) which is used as the column and row indices in the array representation. The x-dimension is the horizontal dimension which increases from left to right and y-dimension is the vertical dimension which will increase downward. The coordinate system for the digital image is depicted in Figure 1.1

1.6 Thesis Organization

Chapter 2 first presents a review on recent and on-going research in automated breast tumor detection from radiographes and introduces the two major approaches in developing automated detection methods, namely *individual tumor type approach* and *mammographic asymmetry approach*. Then the difficulties in developing an automated method for detecting breast tumor from mammogram are discussed. Finally, the overview of processes in mammographic asymmetry approach adopted in this thesis is presented.

In Chapter 3, a detailed discussion on each step of the preprocessing stage of the mammographic asymmetry approach is presented. This includes the mammogram digitization process and breast area segmentation. The control point extraction using a cubic B-spline corner detection method will be discussed and a simple geometric transformation method for aligning the mammogram image pair is proposed. Results on each step to test images are also given.

The design of mammographic asymmetry measurement is discussed in Chapter 4. Four different measurements were investigated which include normalized digital subtraction, locally normalized variance difference, brightness-to-roughness ratio difference and directionality. A mammographic asymmetry measurement which is formed from these four measurements is presented. Also, the criteria for detecting suspicious areas by mammographic asymme-

try is discussed. Results on test images for each of above measurements and final asymmetry measurement are given. Final results of the proposed mammographic asymmetry detection method on ten test cases are presented at end of this chapter.

Chapter 5 will be the conclusion of this thesis and direction on future research in the area of automated breast tumor detection by mammographic asymmetry approach is discussed.

Chapter 2

Review on Automated Breast Tumor Detection

In the past two to three decades, many researchers have demonstrated the potential of using computers in various areas of biomedical image processing. Numerous examples can be found in various scientific journals and research publications. Practical applications have been developed in areas, such as cardiac images enhancement for improved interpretation [MGE+84], boundary detection of ventricle from cineangiograms [CK72], etc. Research directed in the area of automated tumor detection has been particularly intense; and tremendous efforts have been invested, especially in breast tumors detection from mammograms.

Based on the procedures used by expert radiologists, mammographic detection of breast cancers can be classified into two major approaches:

1. *Individual tumor type approach* which is based on detection criteria derived from the set of discriminating mammographic signs of each of the four major types of breast carcinoma described in Section 1.3.2.

2. *Mammographic asymmetry approach* which is based on the fact that human breasts are approximately symmetric and the probability of having simultaneous bilateral breast cancer of the same type at exactly the same corresponding location is extremely low [Urb85]. Any occurrence of breast cancer growth in one breast will produce an asymmetry in identical view mammograms of both breasts. Detection of such asymmetry thus leads to the detection of breast cancer.

The possibility of automating the detection process of breast cancers from digitized mammogram images has been investigated by several groups of researchers. Both of these approaches have been adopted in different automation projects and various degrees of success have been achieved and reported.

Winsberg *et al.* [WEM+67] may have been the first to apply computer image processing technique in detecting breast cancer from mammograms. In their method, digitized images of identical view of both breasts were first subdivided into 64 small 8x8 pixel windows of 16 grey-levels. Each window

is then characterized by 4 feature vectors which measure the distribution of grey-level density within the window. Finally, the vector sets for each breast are compared, resulting in the detection of the presence and location of abnormalities [WEM+67]. Their method is a typical mammographic asymmetry approach method as comparison of both breast images is performed and no specific mammographic feature of any particular breast tumor type is used.

Ackerman and Gose [AG72] developed four measures of malignancy from digitized xeromammograms (calcification, spiculation, roughness, and area-to-parameter ratio) to assign a classification of malignant or benign to tumors. Although, their work was aimed at automated tumor classification and not at detection; they laid the groundwork for the individual tumor type approach of automated breast tumor detection. Their classification criteria which included features such as calcification, spiculation, and roughness (homogeneity) can, in fact, be used as tumor detection criteria.

Hand *et al.* [HSAA79, SSA+80] developed a fully automated system for detecting and locating suspicious areas of breast cancers from digitized xeromammograms. Their method is again based on the mammographic asymmetry approach which compares contralateral identical view of both breasts, using 14 texture parameters and 2 shape parameters defined on 10x10 pixel window. 87% of suspicious areas out of 30 test cases were correctly identified but a very large number of false-positive suspicious areas per case (53.6) were

reported. As described in Section 1.3, xeromammograms have characteristics vastly different from conventional screen-film mammograms. Thus, the effectiveness of Hand *et al.*'s method on screen-film mammogram cannot be determined.

Lai [Lai88] developed an algorithm for detecting suspicious areas of circumscribed tumors from single view mammograms based on the individual tumor type approach. The criteria of her detection algorithm are designed based on the 4 basic mammographic features of circumscribed mass (CM) given by expert radiologists. These mammographic features include the following:

1. CMs have a higher density value (brighter) than surrounding tissue.
2. CMs have a uniform density within the tumor area.
3. CMs have an approximately circular shape of various sizes.
4. CMs have a fuzzy boundary.

Lai's algorithm first uses an edge-preserving smoothing filter, *selective-median filter*, to enhance the digitized mammogram image, followed by a classical template matching process with circular templates to detect circumscribed masses. Furthermore, Lai implemented two additional tests to reduce the number of false-positive suspicious areas from the resulting candidates generated by the template-matching process. A 100% perfect detection rate with an average of 1.7 false-positive detections per case from the 17 test cases was reported.

2.1 Difficulties in Detecting Breast Tumor from Mammograms

Despite all efforts, the success of automated breast tumor detection is very limited. The limited success in all the automated breast tumor detection projects is caused by several factors which can be summarized as follows.

In image pattern recognition, one of the most difficult steps is to extract objects from a background. The degree of difficulty varies greatly with the quality of the picture and the nature of the object [CK72]. The object detection problem has been recognized to be one of the most difficult tasks in biomedical image processing, especially in tumor detection. Like many other types of radiographic images, mammograms possess the same inherent characteristics—low contrast. Obtaining sufficient contrast that would define the anatomical structures of the breast's soft tissue in mammograms is always a major problem. The digitization process not only results in reduction of overall contrast, but also introduces noise, and quantization and sampling errors.

The human breast consists of a number of different anatomical structures such as glandular tissue, blood vessels, ducts, subcutaneous fat, etc. Further, the mammographic appearance of normal breasts varies dramatically between different individuals or for the same individual at different times. This wide variation of mammographic appearance presents a highly irregu-

lar background which makes tumor detection a difficult task.

In order to design an automated method to reliably detect likely tumor sites, the radiologists' description of the X-ray signs of the tumor must be translated into more explicit terms [Lai88]. When interpreting mammograms, expert radiologists use not only the primary mammographic signs described in Section 1.3.2; and, secondary mammographic signs such as asymmetric density or ducts, retraction of skin, nipple or areola; but also their experience in cross-referencing with previous similar cases or the patient's medical history. Hence, the detection criteria given by the expert radiologists are often complex, vaguely defined and difficult to translate into explicit terms. In most research projects, researchers thus use only a subset of mammographic detection criteria which are theoretically unable to cover all the variations of tumor types.

Summing up all the above factors, the design of an automated detection method for breast tumors from mammograms is, indeed, an arduous task. As a result, only limited success in all the automated breast tumor detection projects is achieved.

2.2 Advantages of Asymmetry Approach

There are a few advantages of the mammographic asymmetry approach over the individual tumor type approach in designing automated tumor detection method. First of all, as discussed in the last section, mammographic signs given by expert radiologists are often vaguely described and do not cover all the variations of breast tumors. The mammographic asymmetry approach does not use specific mammographic signs of particular tumor types and thus eliminates errors in translating descriptions of mammographic signs into procedural terms.

Second, detection algorithms for individual tumors frequently tackle one tumor type at a time and the detection criteria often include only a subset of mammographic signs and thus cannot detect all variations of that tumor type. On the other hand, the mammographic asymmetry approach does not depend on the particular mammographic signs of specific tumors. Hence, in principle, it can detect suspicious areas of all types of breast tumors.

Finally, the main objective of a mass screening process is to detect suspicious tumor sites; classification of the suspicious areas into various major categories is of minor concern. Although, the short-coming of individual tumor type approach, detecting one tumor type at a time, can be overcome by integrating various detection algorithms for different tumor types to detect all types of breast tumor; the computation cost will be enormous due

to the redundancy in processing normal breast area by each individual algorithm. The mammographic asymmetry approach can serve as a preprocessing step for further diagnostic procedures such as classification of suspicious areas into benign or malignant nature, or one of the four major categories of breast tumor.

2.3 Processes of Asymmetry Approach

The processes involved in the mammographic asymmetry detection algorithm can be summarized in Figure 2.1. These processes can be generally divided into two major stages, namely, the preprocessing stage and detection stage.

It is common knowledge that human breasts, in nature, are not precisely identical. In addition, the presence of tumor growth in one breast often causes a discrepancy in size and shape between the two breasts. In order to reliably detect mammographic asymmetry caused by tumor growth from a pair of identical view mammogram images, both images must be *perfectly* aligned such that any difference in size and shape must be eliminated as it presents a kind of asymmetry of itself. In the preprocessing stage, a pair of identical view digitized mammogram images are scaled and aligned so that asymmetry measures can be properly performed in the detection stage.

In the detection stage, various asymmetry measures are designed to measure mammographic differences. A decision process is followed to detect

suspicious tumor areas from the above asymmetry measurements. The decision rule is developed from the linear combination of those measurements. In-depth details of each process in the detection algorithm are given in following chapters.

2.4 Examples of Mammogram Images

Plate 2.1 is a typical pair of *Medio-Lateral* view mammograms; whereas, Plate 2.2 is a typical pair of *Cranio-Caudal* view mammograms.

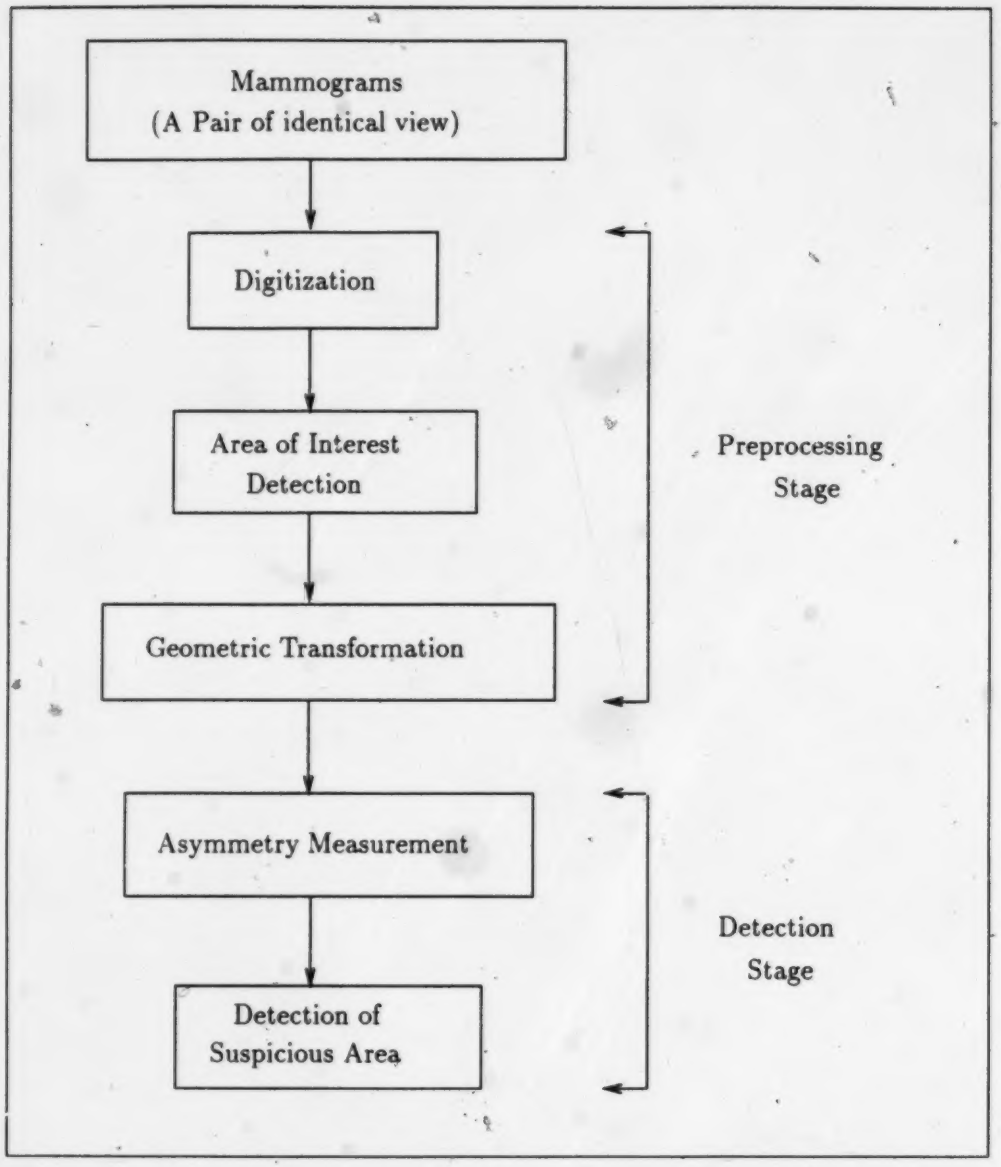


Figure 2.1: Mammographic Asymmetry Approach Processes

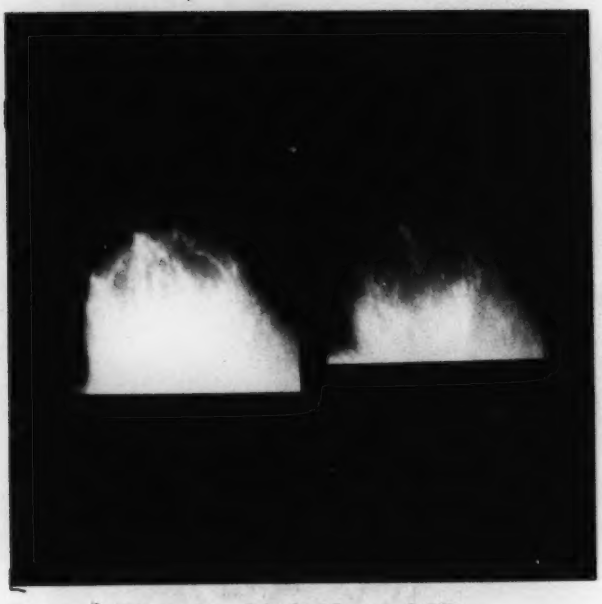


Plate 2.1: Medio-Lateral View

Plate 2.2: Cranio-Caudal View

Chapter 3

Preprocessing of Mammograms

To detect an asymmetry in any comparison process of two objects, a one-to-one mapping of the two objects must be established at the level of comparison. In the detection of mammographic asymmetries from a pair of identical view mammograms, a mapping at an anatomic structural level between two images should be obtained before comparison process can proceed. By obtaining the mapping of the anatomic structures between the two images, natural asymmetries caused by discrepancy in relative position of structures such as glandular tissue, duct, blood vessels within the breast can be eliminated. However, an accurate mapping of anatomic structures is extremely difficult to obtain as there is a wide variation in mammographic appearance of human breast and the detection of various anatomic structures is often complicated. As an alternative, a one-to-one mapping at the pixel level

of both breast images must be obtained. The preprocessing stage of the mammographic asymmetry detection algorithm consists of procedures that digitize and align a pair of mammograms, and eliminate any difference in size and shape so that mammographic asymmetries can be reliably detected in following stages. This chapter presents each step of the preprocessing stage.

3.1 Digitization of Mammograms

Since mammograms are not yet available in digital format, conventional film mammograms must be first converted into digital format. Digitization of film mammograms basically consists of two processes: sampling and quantization. The sampling process is used to extract a discrete set of numbers which represent the brightness values at a regularly spaced array of points. These brightness values are then quantized into a set of equally spaced discrete grey level values.

The quality of the digitized image is solely determined by its resolution and contrast. Spatial resolution of a digitized image in turn depends upon the sampling frequency used in the sampling process; and contrast is determined by the number of discrete grey levels used in the quantization process. These quality measures ultimately depend upon the digitization equipment used.

To avoid overhead computation in rotation and reflection operations in later alignment processes, all mammograms were digitized at the same

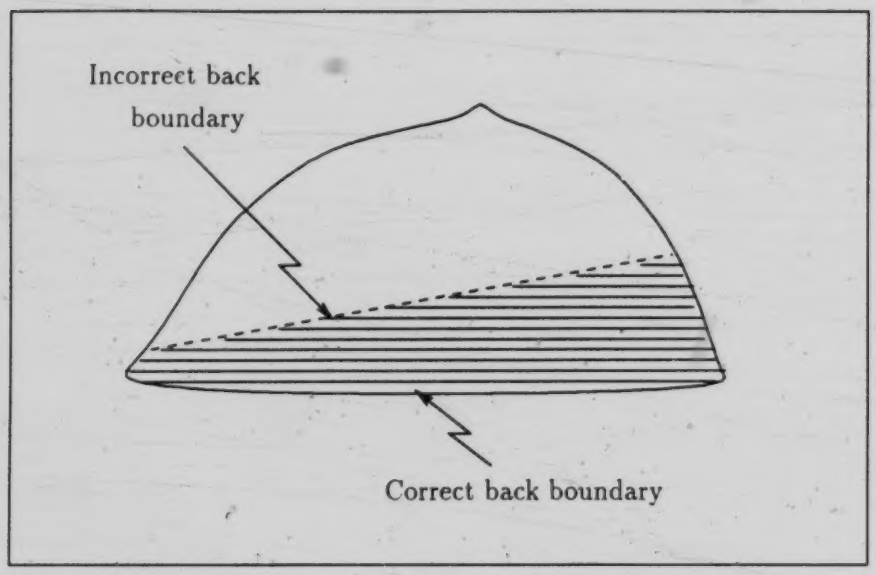


Figure 3.1: Boundary diagram of digitized mammogram

orientation. Furthermore, all lettering such as patient number, date, view, etc. were removed by covering them up with black paper. Most importantly, an accurate back boundary must be provided for each pair of mammograms in the digitization process as the total breast area to be compared is defined by the back boundary. An incorrectly determined back boundary (dotted line in Figure 3.1) eliminates part of breast area (shaded) and will result in incorrect alignment of the image pair, followed by erroneous detection of mammographic asymmetries at a later stage. An example of a digitized image can be found in Plate 2.1 and 2.2.

3.2 Image Enhancement

Like many other types of radiographic images, mammograms are characterized by their low resolution and low contrast ratios. Due to quantization and sampling errors, additional information is lost during the digitization process. Moreover, noise caused by digitization equipment is introduced in digitized images [Lai88]. A large number of image enhancement techniques have been developed to improve the appearance of images by means of contrast enhancement, noise removal, etc. A number of simple and general techniques were studied and their effectiveness on mammogram images were evaluated by Lai [Lai88]. In spite of the fact that some of these enhancement methods improve the observer's ability to analyze images, they do not add information to the image itself. On the contrary, noise removal techniques such as average filtering may remove important features from the image. For these reasons, no enhancement was used in our system.

3.3 Area of Interest Detection

In any comparison process, the objects to be compared must be readily located. By the same token, in order to detect mammographic asymmetry, breast areas must be extracted from the background. It is important that the proper method for extracting breast areas must be used because any

inclusion of non-breast areas or exclusion of breast areas from either one of the images, will lead to erroneous signalling of a mammographic asymmetry. The identification of breast area restricts the area of interest and thus improves the efficiency of subsequent processing.

Many image processing techniques have been developed for the purpose of extracting objects from a picture. Unfortunately, no single method is suitable for all image segmentation problems. Most techniques were designed either for a particular type of application or for segmenting particular kinds of objects. For example, edge-based segmentation algorithms are most suitable for segmenting objects with distinct sharp boundaries; but are inappropriate for images such as mammograms where objects have no well-defined boundaries. Template matching can be used when prior information about the object's shape, size, and orientation is available. Texture-based segmentation techniques are ideal for extracting objects with a texture that has a geometrical or statistical regularity. Thresholding is especially efficient in cases where object and background have grey level values distributed over different ranges.

A proper segmentation technique requires some prior knowledge about the objects and the background of the pictures. In Hand and Semmlow *et al.*'s [HSAA79, SSA+80] study on the detection of suspicious areas in xeromammograms, knowledge about the breast orientation within the image, and size and shape of the nipple is used in designing their segmentation process

of the breast area from the image. In their segmentation process, an edge map is first obtained by thresholding the response of a Sobel filter. Then the most likely position of breast boundary is traced starting with the nipple point. Given a fixed breast orientation, all areas below the breast boundary are eliminated as background. Thus, the breast area is segmented from the image as the back boundary is readily available from their digitization process.

For several reasons the segmentation method used by Hand and Semmlow *et al.* is not suitable for our application. The nipple profile might not appear in a mammogram due to improper positioning or compression of the breast when the mammogram was taken and the image may not reveal the skin boundary of the breast due to the low spatial resolution and the small grey level range.

By analyzing the characteristics of the digitized mammogram images, it is found that the breast is clearly distinguishable from the non-breast background areas. The grey level histogram also shows that the breast area and background areas have grey level values distributed over different ranges and form a bimodal distribution, as shown in the histograms in Figure 3.2. This enables the use of thresholding for extracting the breast area from mammogram images.

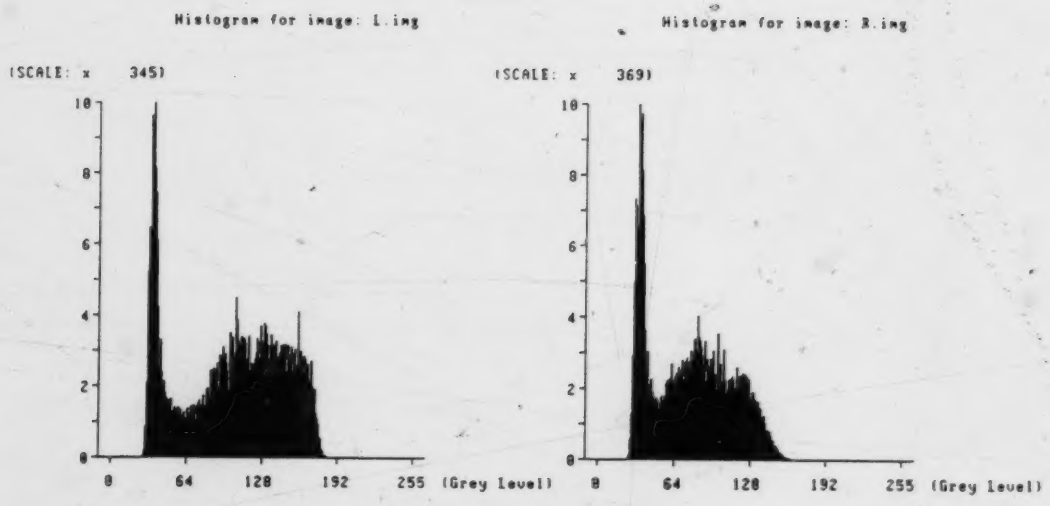


Figure 3.2: Histogram for Plate 2.1

Thresholding an image function $f(x, y)$ at threshold grey level t can be defined as:

$$f_t(x,y) = \begin{cases} b_0 & \text{if } f(x,y) < t \\ b_1 & \text{if } f(x,y) \geq t \end{cases}$$

where b_0 and b_1 are two distinct grey level values. In this project, a simple global thresholding method was used in extracting the breast area from each mammogram image. This simple method is adequate for several reasons. First of all, our digitized mammograms do not reveal the breast boundary very clearly. Therefore, a sophisticated method to detect the breast boundary precisely is not warranted. Furthermore, as discussed in the beginning of this chapter, the comparison process of the mammographic approach should ideally be performed at the anatomic structural level rather than at the

precision of pixel level; thus, minute discrepancies between the detected breast area and the true breast area should not affect the overall accuracy in asymmetry detection.

Threshold selection is done interactively, allowing users to select the threshold value by incrementally and interactively excluding the background pixels with grey levels below the selected threshold value. Due to variations in exposure, the thresholds varied in a range of 35 to 55 grey level in the ten test cases (totalling 20 images). The resulting thresholded image for the left image in Plate 2.1 is shown in Plate 3.1.

As one can observe, the global thresholding method does not yield a breast area with sharp boundary. To overcome this problem, a threshold averaging filter of size 11x11 pixels is used to smooth the binary image. Given an binary image from the above thresholding process, $f(x,y) \in \{0,1\}$ where 0 represents the non-breast background area pixel and 1 represents the breast area pixel, the averaged threshold image is defined as:

$$f_A(x,y) = \begin{cases} 0 & \text{if } \sum_{i=x-5}^{x+5} \sum_{j=y-5}^{y+5} f(i,j) < \frac{11 \times 11}{2} \\ 1 & \text{if } \sum_{i=x-5}^{x+5} \sum_{j=y-5}^{y+5} f(i,j) \geq \frac{11 \times 11}{2} \end{cases}$$

Thus, a pixel is classified as breast area if more than half of the neighbor pixels within the 11x11 pixel window are breast pixels. The result on Plate 3.1 is depicted in Plate 3.2.

3.4 Control Point Extraction

Prior to the geometric transformation operation which matches mammogram image pairs, some visible and unambiguously defined features must be detected and matched correspondingly. These unique features play a very important role as control points in the image registration process. Goshtasby points out that two images that have translational, rotational, and scaling differences can be registered if the coordinates of at least two pairs of corresponding control points from the two images are known [Gos83]. Also, Zhou *et al.* [ZG88] have suggested that if a sufficient number of corresponding features or "control points" in the mammograms could be determined, then mammogram images can be accurately registered over their whole area.

Since the comparison process should be performed at an anatomic structural level; our mammographic asymmetry detection algorithm does not demand a geometric transformation which matches the images precisely. Additionally, the complexity and amount of human assistance in selecting a sufficient number of control points for a sophisticated geometric transformation algorithm is prohibitive for a mass screening procedure. Thus, a simple geometric transformation algorithm requiring a limited number of control points appears to be adequate for this project.

3.4.1 Difficulty in Automating Extraction

Due to the large variation in mammographic appearance of the human breast, it is difficult to identify unique features automatically within the breast area. Therefore, it is arduous, if not impossible, to develop automated methods that can accurately determine corresponding control points within the breast area. In the geometric unwarping algorithm for digital subtraction mammography by Zhou *et al.* [ZG88], control points were manually selected instead of automatically detected with a computer program.

3.4.2 Proposed Approach

Even if control points can not be easily detected within the breast area, some control points can be easily obtained from the breast boundary, namely the nipple and the intersection of the breast boundary and the back boundary as depicted in Figure 3.3. These control points are corners on a smooth curve boundary with each control point having a higher curvature compared to the rest of the boundary. Various corner detectors have been developed and studied in the field of image processing. Medioni and Yasumoto [MY87] have studied a number of corner detectors and found that their proposed cubic B-spline method is excellent in terms of computation cost and accuracy. Hence, their cubic B-spline method is adopted in this project.

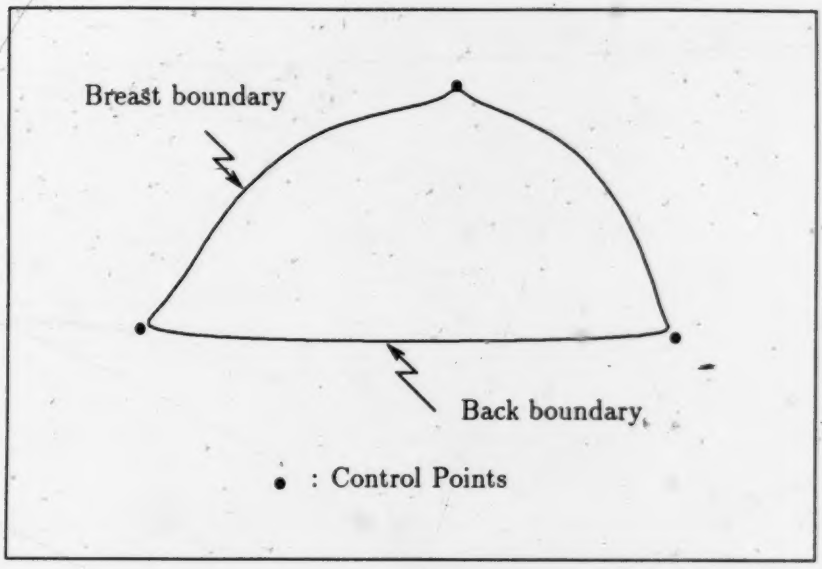


Figure 3.3: Control points of a typical breast area

The overall approach for extracting the control points for the geometric transformation operation includes the following steps:

1. Breast boundary extraction and smoothing.
2. Control point extraction by cubic B-spline corner detection method.

As discussed in section 3.3, the nipple profile might not appear in the digitized mammogram image. An interactive program is written to allow the user to enter the nipple position. Although this requires the user's assistance in the process; the amount of assistance is limited. Furthermore, the nipple profile should appear in the digitized mammogram, if the mammogram was taken with proper position, compression and exposure settings.

3.4.3 Breast Boundary Extraction and Smoothing

Before the cubic B-spline method can be used, the boundary pixels of the breast area must be extracted. A simple edge tracing method is used to extract the boundary coordinates of the breast area from the binary image obtained in the previous area of interest detection process. The resulting coordinates are stored in an array data structure.

Due to the nature of global thresholding technique, the resulting breast area detected in the previous step has a rather rough boundary; therefore, a one-dimensional averaging filter is used to remove this excess jitter. Mathematically, given an array of coordinates

$$B = \{P_i = (x_i, y_i) \mid 0 \leq i < n; i + n = i\}$$

which represents a closed curve of length n pixels, the one-dimensional average filtering on B with a filter size N will yield

$$B' = \{P'_i = (x'_i, y'_i) \mid 0 \leq i < n; i + n = i\}$$

where:

$$x'_i = \frac{1}{N} \sum_{j=i-\lfloor N/2 \rfloor}^{i+\lceil N/2 \rceil - 1} x_j$$

$$y'_i = \frac{1}{N} \sum_{j=i-\lfloor N/2 \rfloor}^{i+\lceil N/2 \rceil - 1} y_j$$

A filter size of $N = 31$ was selected as it yielded satisfactory results for all test cases. The smoothed breast boundary detected for the left image in Plate 2.1 is depicted in Plate 3.3.

3.4.4 Cubic B-Spline Method

As stated in [MY87], a cubic polynomial is used to fit a curve between point A and B with parameter t varying between 0 and 1 as follow:

$$\begin{aligned} x = f(t) &= a_1 t^3 + b_1 t^2 + c_1 t + d_1 \\ y = g(t) &= a_2 t^3 + b_2 t^2 + c_2 t + d_2 \end{aligned} \tag{3.1}$$

The curvature at point $A(t = 0)$ can be derived as:

$$C_v(0) = 2\frac{c_1 b_2 - c_2 b_1}{(c_1^2 + c_2^2)^{3/2}} \tag{3.2}$$

Suppose $P_{i-1}, P_i, P_{i+1}, P_{i+2}$ are equally spaced neighboring points on a given curve as shown in Figure 3.4 and the approximation of the segment between P_i to P_{i+1} by a cubic B-spline is given:

$$\begin{aligned} x(t) &= TMP_x \\ y(t) &= TMP_y \end{aligned} \tag{3.3}$$

where:

$$\begin{aligned} T &= (t^3, t^2, t, 1) \\ M &= \frac{1}{6}\begin{bmatrix} -1 & 3 & -3 & 1 \\ 3 & -6 & 3 & 0 \\ -3 & 0 & 3 & 0 \\ 1 & 4 & 1 & 0 \end{bmatrix} \end{aligned}$$

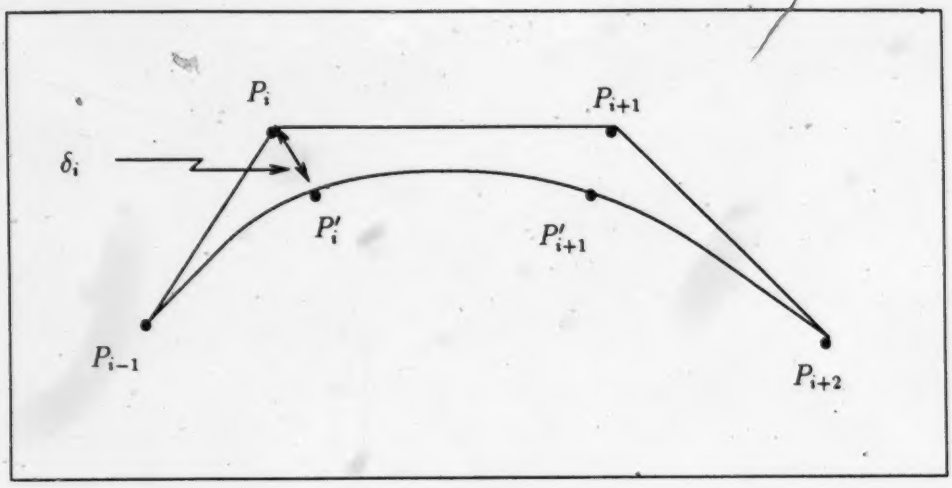


Figure 3.4: Cubic B-spline approximation for P_i to P_{i+1}

$$P_x = \begin{bmatrix} x_{i-1} \\ x_i \\ x_{i+1} \\ x_{i+2} \end{bmatrix} \quad ; \quad P_y = \begin{bmatrix} y_{i-1} \\ y_i \\ y_{i+1} \\ y_{i+2} \end{bmatrix}$$

Expanding equation 3.3 to the form of equation 3.1, the curvature at P_i can be computed in terms of the coordinates of P_{i-1}, P_i, P_{i+1}, and the displacement between the original position of P_i and the point on the interpolating spline P'_i is given as:

$$\begin{aligned} \delta_{ix} &= x_{i-1}/6 - x_i/3 + x_{i+1}/6 \\ \delta_{iy} &= y_{i-1}/6 - y_i/3 + y_{i+1}/6 \end{aligned} \tag{3.4}$$

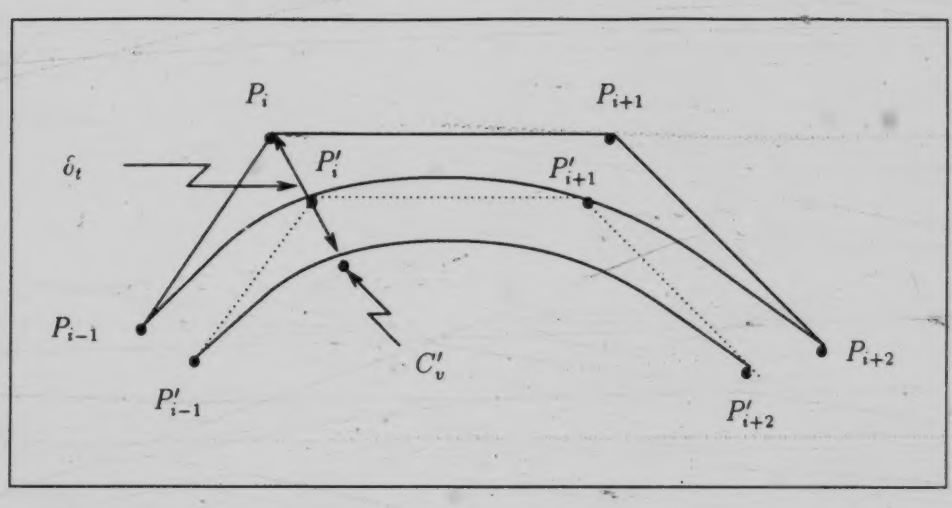


Figure 3.5: Cubic B-spline approximation for P'_i to P'_{i+1}

The first interpolation serves as a smoothing process to reduce the effects of quantization on slope or curvature [MY87]. A second interpolation is performed on the displacement points $\{P'_j = P_j + \delta_j \mid i-1 \leq j \leq i+1\}$ to detect corners. The second interpolation process is depicted in Figure 3.5. The curvature C'_v for the point P_i of the second interpolation can be computed by equation 3.2 where:

$$
\begin{aligned}
b_1 &= \frac{x_{i-2}}{12} + \frac{x_{i-1}}{6} - \frac{x_i}{2} + \frac{x_{i+1}}{6} + \frac{x_{i+2}}{12} \\
b_2 &= \frac{y_{i-2}}{12} + \frac{y_{i-1}}{6} - \frac{y_i}{2} + \frac{y_{i+1}}{6} + \frac{y_{i+2}}{12} \\
c_1 &= \frac{(x_{i+1} - x_{i-1})}{3} + \frac{(x_{i+2} - x_{i-2})}{12} \\
c_2 &= \frac{(y_{i+1} - y_{i-1})}{3} + \frac{(y_{i+2} - y_{i-2})}{12}
\end{aligned}
\tag{3.5}
$$

and total displacement from the original position P_i is δ_t where:

$$\delta_{tx} = x_{i-2}/36 + 2x_{i-1}/9 - x_i/2 + 2x_{i+1}/9 + x_{i+2}/36$$

$$\delta_{ty} = y_{i-2}/36 + 2y_{i-1}/9 - y_i/2 + 2y_{i+1}/9 + y_{i+2}/36$$

The criteria used by Medioni and Yasumoto to classify a point a corner are:

1. the total displacement δ_t is greater than a given threshold value d_c.
2. the curvature C'_v is greater than a given threshold C_c.
3. the curvature C'_v is a local maximum.

This set of criteria can be simplified in the control point extraction process since only a fixed number of corners (3) is required to be detected and these points should have the maximum curvature among all points of the breast boundary. Therefore, the three local maxima with maximum curvature are selected.

In the implementation of this method, an interval of 8 pixels is used to select neighboring points for each pixel along the breast boundary detected earlier. The curvature for each point is then computed using equation 3.2 with coefficient formulas given in equations 3.5. To locate the local maximum of curvature accurately, a one-dimensional Gaussian filter is used to smooth the curvature results. The Gaussian filter is defined mathematically as:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp -x^2/2\sigma^2$$

where σ is a parameter controlling the amount of smoothing. To generate the Gaussian filter, $\sigma = 1.0$ is used. Three local maximum curvature points with highest curvature value are selected as control points after smoothing with the above filter. An example of control points detected for the left image in Plate 2.1 is depicted in Plate 3.4.

3.5 Geometric Transformation

As mentioned earlier, natural mammographic asymmetries exist in any given pair of identical view mammograms. This is caused not only by the differences in size and shape between the left and right breasts but also by the differences in patient positioning and compression settings when the mammogram was taken. This type of asymmetry must be eliminated in order to accurately detect only mammographic asymmetry that was caused by the growth of breast carcinoma. The geometric transformation process aims at minimizing the amount of natural mammographic differences which is not caused by breast carcinoma.

3.5.1 Difficulties In Obtaining Accurate Method

It is difficult to obtain an accurate geometric transformation which aligns a pair of identical view mammograms and yields minimum natural asymmetry. This is not only due to the aforementioned factors, but also due to

deformation during the breast compression in the mammography procedure. The human breast is a 3-dimensional flexible soft-tissue organ and conventional film-mammography is only a 2-dimensional projection image of this 3-dimension organ; hence, the amount of shearing effect from breast compression is not available in these 2-dimensional mammograms. The shearing effect in one direction on the projection plan is depicted in Figure 3.6. Note that the shearing can occur in any direction that is perpendicular to the breast's thickness dimension.

3.5.2 Review on Existing Methods

Although many techniques have been developed for image registration, most methods are inadequate for mammographic images. They are either too simple and do not yield adequate results or too complicated to be used effectively and feasibly in a mass screening program. For example, Hand and Semmlow *et al.*'s method [HSAA79, SSA+80] is based on a simple 2-dimensional translation function which minimizes the least square error between the two breast boundaries. Their method is simple and easy to implement; however, it is not adequate for mammographic asymmetry detection. This is because the lack of geometric transformation to eliminate the natural discrepancy in size and shape between both breasts, therefore causing this discrepancy to be falsely classified as a mammographic asymmetry.

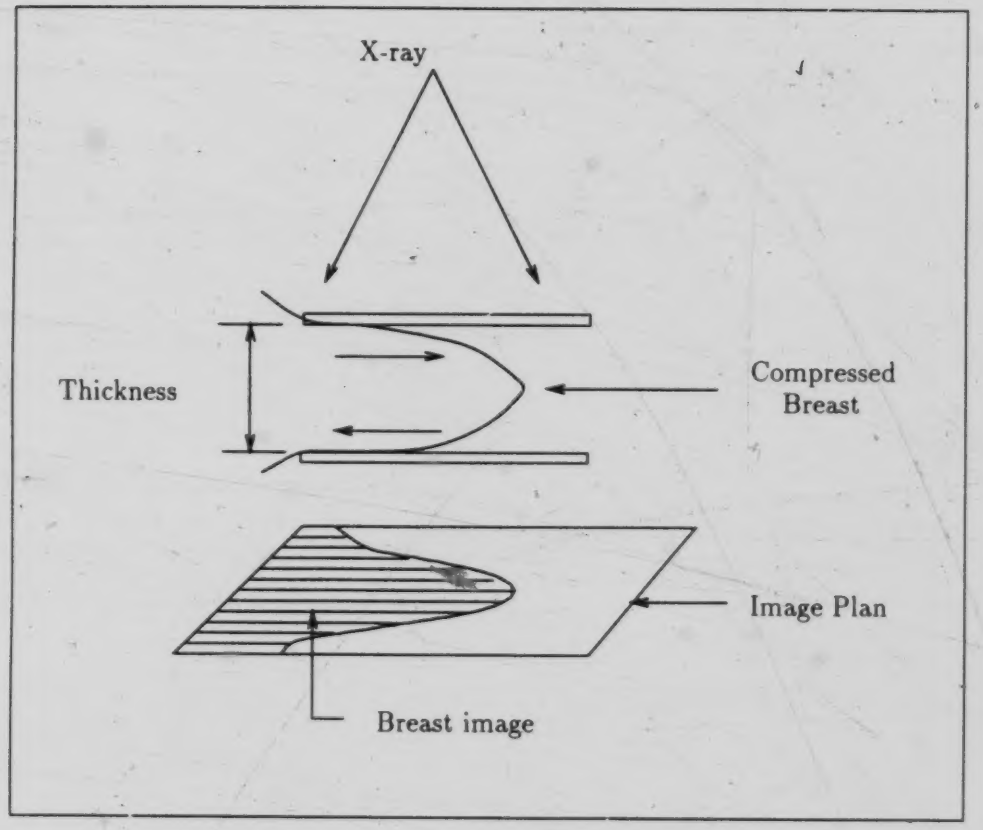


Figure 3.6: Shearing effect of breast compression

On the other hand, Zhou and Gordon [ZG88] employ a rather sophisticated but complicated technique in unwarping mammogram images for digital subtraction. Firstly, their method requires the input of a sufficient number of control points within the breast area, then a triangulation process partitions the convex hull of this set of control points into triangles. After estimating the partial derivatives of each control point, separate interpolation functions are derived for each triangle to compute the interpolated coordinate for each point within its respective triangle. Clearly, this local transformation method is more elaborate than the one used by Hand and Semmlow *et al.*; however, the amount of human assistance in selecting control points and the high computation cost, especially in the two-dimensional point-inclusion problem, make it infeasible to be used in a mass screening procedure. Hence, a simple geometric transformation for the mammographic asymmetry detection algorithm has to be found.

3.5.3 Proposed Method

A 2-dimensional geometric transformation can be mathematically defined by a pair of equations in the form:

$$x' = f_1(x, y) \qquad , \qquad y' = f_2(x, y)$$

which specify the new coordinates (x', y') of each point of the transformed image as a function of the old coordinates (x, y). Geometric transformations

are commonly classified into two categories, *global* and *local* techniques. In global geometric transformation one set of transformation functions $\{f_1, f_2\}$ is used to transform the entire image; whereas in local geometric transformation the image is first subdivided into subimages and separate transformation functions are used to transform each subimage.

Zhou and Gordon [ZG88] suggested that geometric distortions in many images are due to local factors and they have demonstrated that local geometric transformation method, based on a collection of local transformation functions, is more accurate in unwarping mammogram images for digital subtraction. However, the programming complexity and computational cost is also higher than it is for the global method, and is in directly proportion to the number of subimages that the geometric transformation method subdivided from the image. Hence, there is a tradeoff between accuracy and computational cost in term of the degree of locality of the transformation method. In the following, a local geometric transformation method is presented that has sufficient degree of locality to yield satisfactory results in aligning a pair of identical view mammograms.

For matching a pair of objects, one can either transform both objects into a common reference model for comparison or use one of the objects as a reference model and transform the other object to match that reference object. Obviously, the latter approach is better than the first one for several reasons. First of all, resampling or interpolation errors are introduced into

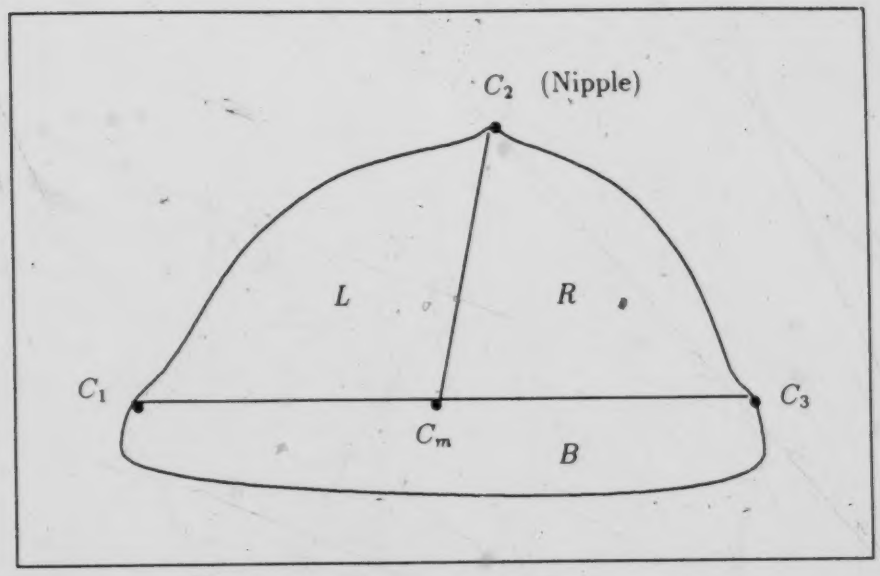


Figure 3.7: Subdivision of breast area

only one image [RK82, section 9.3.1]. Second, the complexity and computational cost is lower. Therefore, the mammogram with the larger breast area from a given pair of mammogram images is selected as the reference image for transformation. This choice of selection eliminates the problems of lost information and resolution reduction that are involved in transforming a larger breast area into a smaller one.

Once the reference image is identified, the breast area in both images is subdivided into three regions based on the control points detected in the previous section. The subdivision is depicted in Figure 3.7. To facilitate subsequent transformation procedures, both images are first rotated such

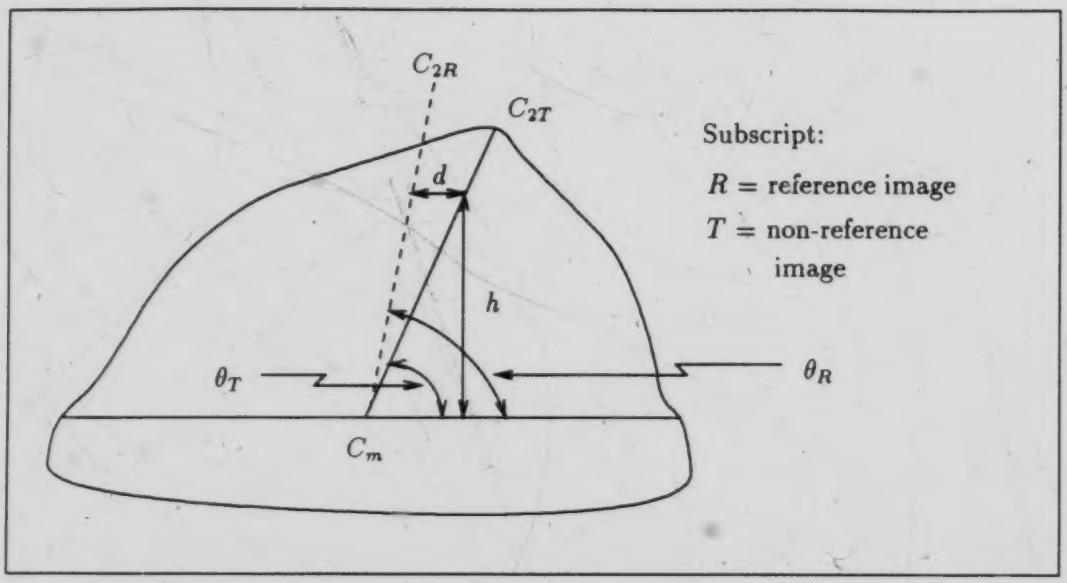


Figure 3.8: Elimination of skewing effect

that their baselines $\overline{C_1C_3}$ are parallel to the x-axis of the coordinate system and then the non-reference image is translated such that the mid-points of both baselines are aligned.

The next step is to eliminate the skewing effect above the baseline that is caused by breast compression (see Figure 3.8). Although the skewing effect is two dimensional, only the horizontal skewing effect is considered at this stage. For every row of pixels above the baseline of the non-reference image, they are shifted horizontally by:

$$d = h \times \left(\frac{1}{\tan \theta_R} - \frac{1}{\tan \theta_T} \right)$$

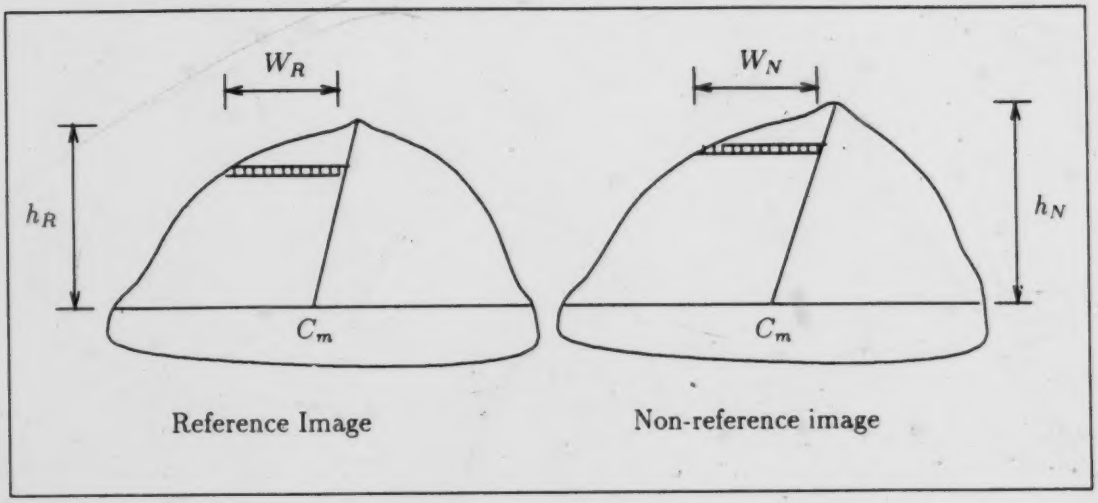


Figure 3.9: Transformation of sub-region L

where: h = the distance in y-dimension between the row of pixel to be shifted and the baseline.

A common simple bilinear interpolation algorithm is used to transform the sub-regions L and R depicted in Fig 3.7. An example of transforming the L sub-region is shown in Figure 3.9. First, the algorithm finds the ratio,

$$H_ratio = \frac{h_N}{h_R} \ ,$$

i.e. the ratio of the baseline-to-boundary distances in both breast images. This ratio maps every row of pixels in the reference image to the non-reference image. Then, a ratio of the widths of the corresponding row of pixels between

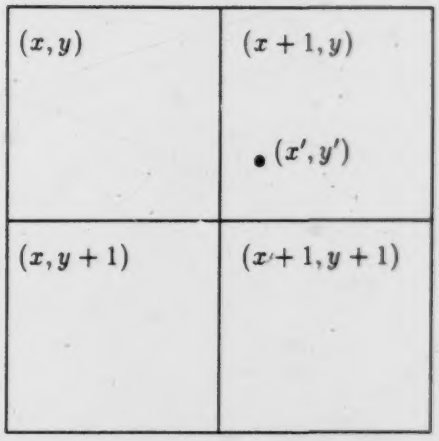


Figure 3.10: Bilinear transformation of non-reference image

these two images can be formed as:

$$W_ratio = \frac{W_N}{W_R}$$

Thus, every pixel in the L and R sub-regions of the reference image can be mapped to a coordinate of the corresponding sub-region of the non-reference image using these two ratios. The mapped coordinate (x', y') in the non-reference image may not necessarily be integers (see Figure 3.10). A bi-linear interpolation method is then used to assign a grey level value to the transformed image of the non-reference image as follows: let x and y be the integer part of the mapped coordinate x' an y' such that

$$\begin{aligned} x &= \lfloor x' \rfloor \\ y &= \lfloor y' \rfloor \end{aligned}$$

Also, let the fractional part of x' and y' be α and β; i.e.,

$$\alpha = x' - x$$

$$\beta = y' - y$$

Then, the grey level value of the corresponding pixel of the mapped coordinate is given by [RK82, section 9.3.2]:

$$(1-\alpha)(1-\beta)f(x,y)+(1-\alpha)\beta f(x,y+1)+\alpha(1-\beta)f(x+1,y)+\alpha\beta f(x+1,y+1)$$

The deformation of sub-region B in Figure 3.7 is negligible, in particular in the vertical direction. Therefore, we can use a simple horizontal scaling to transform sub-region B. The results of geometric transformation of the pair of images in Plate 2.1 are depicted in Plate 3.5. The top half shows the original image pair and the bottom half is the corresponding transformed images. This shows that the geometric transformation method presented in this section yields satisfactory results.

Plate 3.1: Binary image for the left image of Plate 2.1

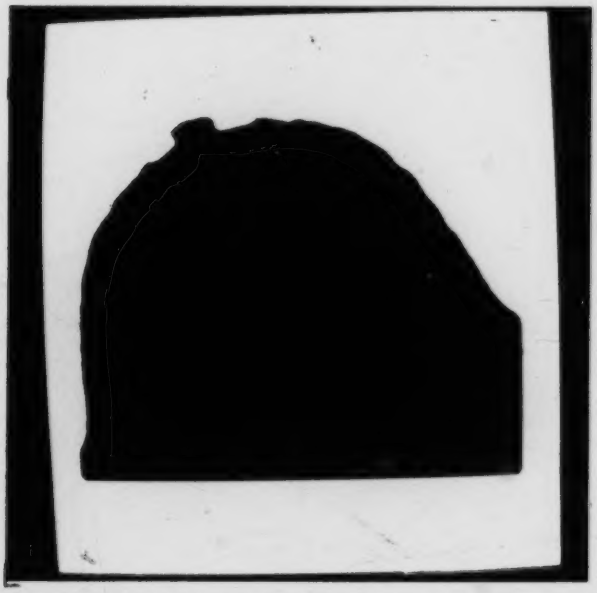


Plate 3.2: Thresholded averaging of binary image in Plate 3.1

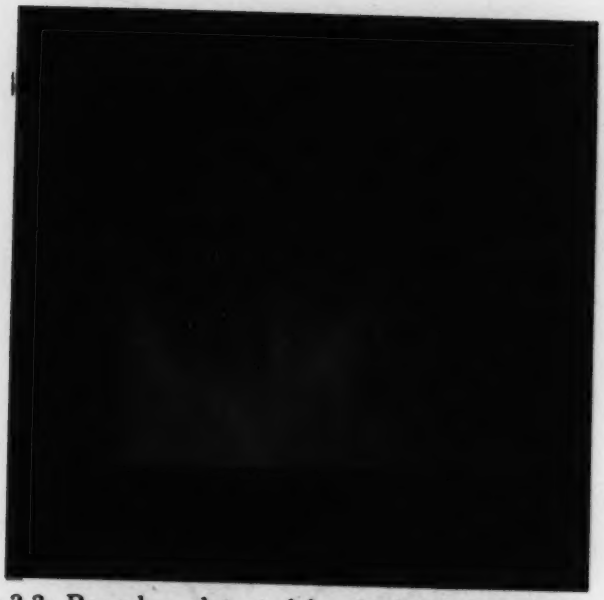


Plate 3.3: Boundary detected for the left image of Plate 2.1

Plate 3.4: Control points detected for Plate 3.3

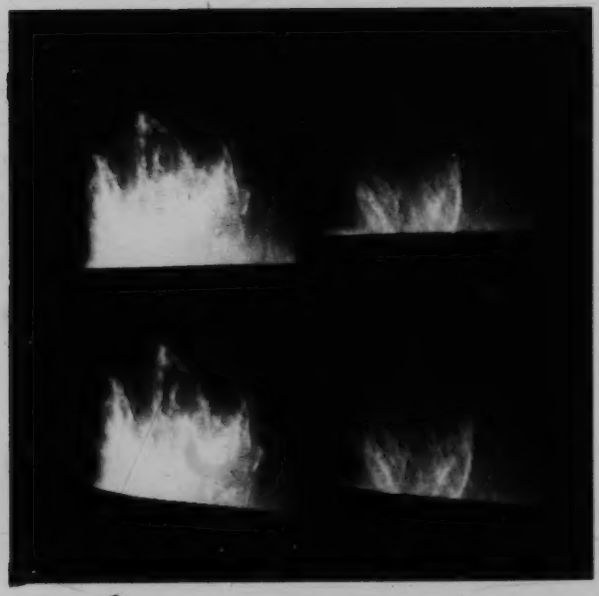


Plate 3.5: Geometrically transformed images for Plate 2.1

Chapter 4

Detection of Mammographic Asymmetry

Any comparison process requires the definition of a comparison measure. For instance, *meter* is used as a measuring unit in comparing the length of two objects; while *degree* is used in comparing two angles. There is no well-defined measurement method available for describing the mammographic appearance of a human breast. This chapter describes some criteria that can be used for comparison and describes how these criteria can be computed.

4.1 Digital Subtraction

Digital subtraction is a basic operation for detecting differences between a pair of digitized images. Zhou and Gordon [ZG88] suggested that pairs of aligned mammograms could be subtracted to highlight changes due to a growing breast carcinoma. By the same argument, one can also detect mammographic asymmetry from a pair of identical view mammograms using digital subtraction.

Since cancer cells have a higher density than normal breast tissue; breast carcinomas appear as structures with higher intensity level (brighter) than normal breast tissue in mammograms. Thus, large discrepancies in grey-level value between two identical view mammograms at any pixel position could indicate breast cancer growth. However, due to the difference in exposure and digitization setting, natural discrepancies in overall brightness and contrast between the two images are unavoidable. These discrepancies in overall brightness and contrast could then be mistaken as mammographic asymmetries caused by cancer growth. To avoid this, a normalization process is performed to eliminate any difference in overall brightness and contrast of the two images. The grey-level normalization process of a given grey-level image, $f(x, y)$ is given by:

$$f_n(x, y) = \frac{f(x, y) - \mu}{\sigma}$$

where μ is the mean grey-level value and σ is the standard deviation of the

entire breast area in the image.

In addition, our geometric transformation does not yield a pixel-precision matching between two mammogram images and sampling errors might be introduced into the transformed image during the geometric transformation process. Therefore, instead of simply subtracting corresponding pixels, the mean grey-level value of neighboring pixels, within a certain window size, is used in the digital subtraction process. Given the normalized reference image, $f_n(x,y)$, and normalized transformed image, $g_n(x,y)$, the digital subtraction process is given by:

$$S(x,y) = \left| f_n(x,y) - \frac{1}{n \times n} \sum_{i=x-\lfloor \frac{n}{2} \rfloor}^{x+\lceil \frac{n}{2} \rceil - 1} \sum_{j=y-\lfloor \frac{n}{2} \rfloor}^{y+\lceil \frac{n}{2} \rceil - 1} g_n(i,j) \right|$$

In other words, the normalized transformed image is first convolved with a averaging filter of size nxn before digital subtraction takes place. Based on visual judgment on the results, a window size of 11x11 pixels was used in the implementation and the result for the images in Plate 2.1 is shown in Plate 4.1.

Digital subtraction is a good measurement for detecting mammographic asymmetry if both images are aligned perfectly. However, perfect alignment of a pair of mammographic images is difficult, if not impossible, to obtain without human assistance and at a high cost. It is very effective for mammograms that do not have much grey-level variation due to anatomic structures such as glandular tissue, ducts, etc., as is the case with the mammograms

in Plate 4.2. This plate shows the geometrically transformed original images on the top half and the result of digital subtraction, after linear grey-level stretching has been applied, is shown at the bottom left. On the other hand, digital subtraction is quite susceptible to noise in images and does not produce a strong response at tumor locations for mammograms with large grey-level variation or for mammograms in which a tumor is partially or entirely hidden in other anatomical structures. Hence, detection of mammographic asymmetry cannot solely depend on this single measurement.

4.2 Analysis Using Texture Statistics

At present, there is no single qualitative measurement that can describe mammographic appearance adequately such that mammographic asymmetry can be determined reliably. However, various anatomic structures of the human breast do appear with different texture characteristics in mammogram images and these characteristics can be represented by a collection of texture measurements developed in texture analysis. Each measurement can be used to reflect a single characteristic difference in a pair of mammogram images and a sufficient number of carefully selected measurements will yield an acceptable level of accuracy in detecting mammographic asymmetry. Hence, detecting mammographic asymmetry can be formed as a problem of texture analysis of a given pair of mammogram images.

Digital texture analysis has been used in image processing applications such as segmentation in scene analysis, object's shape determination, vegetation type classification in LANDSAT images, and evaluation of biomedical images, etc. Despite the importance and wide applicability of texture analysis in image processing, a precise definition for the notion of texture does not exist. However, texture is commonly viewed as a structure composed of either deterministically (regularly or structurally) or stochastically (irregularly or randomly) interwoven elements; and it has the characterizing shift-invariance property, that is, visual perception is independent of the position in the image pattern [Lev85, ch.9]. These "interwoven elements" can be characterized by some mathematical feature descriptions as stated by Haralick [Har79]:

> Image texture can be qualitatively evaluated as having one or more of the properties of fineness, coarseness, smoothness, granulation, randomness, lineation or being mottled, irregular or hummocky. (p.796)

In texture analysis, the first and most important task is to extract texture feature descriptions which embody information about the spatial distribution of intensity variations in the original image [HWG87]. There are two major approaches in texture feature extraction, namely the *structural* and *statistical* approach. The structural approach regards texture as being composed of primitives (interwoven elements) that form a repeating pattern

and describes such patterns with a grammar for generating them [BB82, p.163]. This approach is best for describing texture composed of deterministically placed primitive elements such as "reptile" and "wire braid" textures. Nevertheless, not all textures have a patterned geometric regularity; on the contrary, most textures exhibit patterns of stochastically distributed primitive elements. These highly irregular and random texture patterns are best analyzed by the statistical approach. The statistical approach describes texture patterns by a *feature vector*. The feature vector is a set of statistical measurements which condenses the description of relevant properties such as fineness, coarseness, smoothness, etc. of the textured image into a small feature space for analysis [BB82, p.181]. The statistical approach is most widely used for computer vision applications.

The choice between the two approaches is highly dependent on the nature of the application and the characteristics of the textured images. For mammographic images, the structural approach is not appropriate since the human breast does not reveal any geometric regularity of repeating patterns. In fact, the wide variation of mammographic appearance shows a high degree of irregularity or randomness in grey-level distribution of pixels. Thus, the statistical approach of the texture analysis appears to be the best choice in describing mammographic patterns and detecting mammographic asymmetry.

Within the statistical domain of texture feature extraction, there are many methods for analyzing texture characteristics. These include an auto-correlation function for characterizing texture coarseness, two-dimensional Fourier transform for analyzing texture in spatial-frequency domain, as well as simple first-order statistical measurements such as mean, variance, and skewness of grey-level distribution. However, grey-level co-occurrence matrices and second-order statistics derived from co-occurrence matrices are widely used in defining texture features. Haralick *et al.* [HSD73] defined a set of texture features including angular second moment, contrast, correlation, etc.; and Connors *et al.* [CTH84] defined measures such as inertia, local homogeneity, energy, etc. Nevertheless, the large amount of computation time involved in computing second-order statistics from co-occurrence matrices on large images represents a severe drawback and receives a lot of criticism in the literature. A detailed review of the above techniques can be found in [Lev85, ch.9].

Each statistical measurement can only reflect a single texture property. Variance, for example, reflects the coarseness of a texture. In order to precisely represent a given texture, a large number of statistical feature measurements must be used. The selection of proper texture features is important since it is desirable to have a minimum number of features that can adequately represent the properties of a given texture.

In the past, researchers have used various sets of statistical measurements to represent the properties of mammographic pattern for the detection of breast abnormalities, the classification of breast lesion and the detection of mammographic asymmetry caused by the growth of breast carcinoma. [WEM+67, KOS77, HSAA79, SSA+80]. Hand *et al.* [HSAA79] used 14 texture features constructed from a 10x10 pixel window of the original image data. These texture features include average intensity, 8 roughness measures defined as running sums of absolute intensity change between pixels in both vertical and horizontal directions, and 5 directionality measures which form a running sum of change in intensity across and down the boundaries of the 10x10 pixel window. In addition, 2 shape parameters, circle-likeness and star-likeness, are also used in their detection algorithm.

In [HSAA79], a total of 16 feature parameters were used to detect suspicious areas in breast images. The number of feature parameters can be reduced by improving some of these feature measurements. In fact, a reduced and improved set of feature parameters was used in a subsequent study [SSA+80] where only 6 feature parameters were used to detect suspicious areas. They include average intensity, circle-likeness, star-likeness; and a normalized variance measure was used to replace the 8 roughness parameters in the previous study and 2 directionality parameters were used instead of 5 in the previous study. However, the effectiveness of this improved and reduced feature parameter set is unknown as the suspicious areas detected

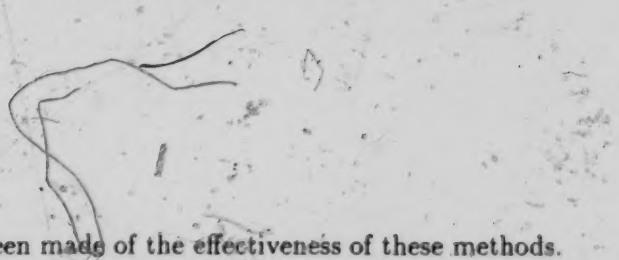
were passed to a breast tumor classification program and no intermediate result on suspicious area detection was presented in their paper [SSA+80].

If our asymmetry detection system is to be used in a mass screening program, we must obtain a minimum set of texture features which is inexpensive to compute and yields a high level of detection accuracy. Several texture measurements were investigated and implemented. The details of the implementation and results are presented in the following sections.

4.3 Roughness (Variance)

Tumor growth often causes changes of mammographic appearance in terms of roughness. For example, the center of a typical stellate lesion would have a dotted pattern due to the longitudinal projection of spicules [AG72]; whereas a typical circumscribed lesion would have a smooth texture compared to the normal breast tissue. Thus, roughness is another measurement for detecting mammographic asymmetry caused by breast tumor growth.

Roughness is a commonly used property in texture analysis. There are many methods for measuring roughness such as, for example, the autocorrelation function or variance measure. In addition, roughness can also be analyzed in the spatial-frequency domain by examining the distribution in the power spectrum of the image or it can be defined using the grey-level co-occurrence matrix [Lev85, ch.9]. However, no direct comparisons have



been made of the effectiveness of these methods.

In breast radiography analysis, several roughness measures have been used. Ackerman and Gose [AG72] defined roughness as the sum of the absolute differences between points with a distance of 2 to 16 pixels along a vertical line. The ratio between the roughness measures of a center circle and an annulus was used as one of the parameters for classifying breast lesions in xeromammograms. Following the same idea, Hand *et al.* defined 8 vertical and horizontal roughness measures for their detection algorithm using the mammographic asymmetry approach [HSAA79]. Since the computation of the above 8 roughness measures is quite costly, Semmlow *et al.* improved the efficiency of their detection algorithm by replacing them with a single normalized variance measure [SSA+80]. Given its efficiency, the variance measure seems to be a good choice for roughness measurement and is adopted in our detection algorithm.

4.3.1 Globally Normalized Variance

Variance is a first-order statistical measure derived from the grey-level histogram and provides an indication of how uniform or regular an image region is. The variance of a region R with an area of n pixels in an image, $f(x, y)$, can be computed by the formula

$$\sigma_R^2 = \frac{1}{n} \sum_{i,j \in R} [f(i,j) - \mu]^2 \quad ,$$

where μ is the mean grey-level value of the region R. This equation can be expanded into the following form

$$\sigma_R^2 = \frac{n\sum_{i,j\in R}[f(i,j)]^2 - \left[\sum_{i,j\in R} f(i,j)\right]^2}{n^2}$$

and this form is used in our implementation. Using a ring-buffer to save the sum and sum of squares in each window, the amount of computation is reduced in subsequent windows.

Once again, the overall brightness and contrast difference between the two images must be eliminated first. Therefore, globally normalized images from the previous digital subtraction process were used. Moving windows with size ranging from 25 to 51 were used in computing the absolute difference of variance for each pixel. The results indicate that the response of this measurement is dominated in an area where a large amount of variation in brightness exists. On the other hand, in an area with low variation in grey-level, the response is weak. A good example for such situation is depicted in Plate 4.3 which shows the original transformed images, Plate 2.1, in the top half, and response of the absolute difference in variance in the bottom-left of the image. The red circle indicates where radiologists suspect the presence of a tumor. As one can observe, this measurement produces a strong response in the left half of the breast area where there are a lot of anatomic structures such as glandular tissue and ducts; whereas the suspicious area is located in a low grey-level variation of fatty tissue background and it only shows a

very weak response to the absolute difference of variance in that area. The reason for such a result is quite logical as the range of absolute difference of variance for areas with high grey-level variation is comparatively wider than the one for areas with low grey-level variation. However, it is desirable to have response for the same amount of absolute variance difference be emphasized in a low grey-level variation neighborhood than in a high grey-level variation neighborhood. A modified roughness measure is proposed in the next section.

4.3.2 Locally Normalized Variance

Due to the linearity characteristics of global normalization, the response of absolute variance difference is dominated in the high grey-level variation area. To overcome this problem, a local normalization approach must be adopted for the roughness measurement.

This local normalization approach operates on two concentric square windows as illustrated in Figure 4.1. The inner window R is used for measuring the absolute difference of variance and the outer window N is used for normalization. Let μ_N, σ_N^2 be the mean and variance of the normalization window respectively. The local normalization of each pixel, $f(x,y)$, in variance window R is given by:

$$f_N(x,y) = \frac{f(x,y) - \mu_N}{\sigma_N}$$

Then, the roughness measure is defined as the variance of locally normalized

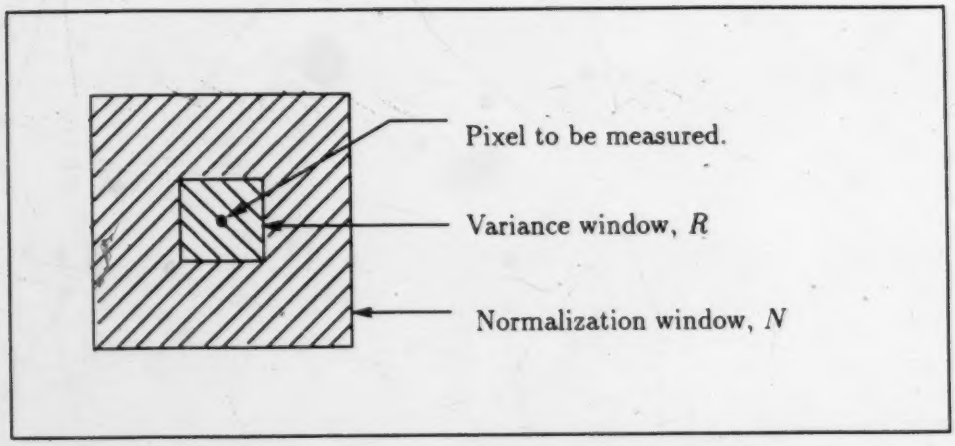


Figure 4.1: Window subdivision for locally normalized variance.

variance window R, such that for $(x, y) \in R$,

$$\begin{aligned}
\textbf{Roughness} &= \textbf{Variance}\,[f_N(x,y)] \\
&= \textbf{Variance}\left[\frac{f(x,y)-\mu_N}{\sigma_N}\right] \\
&= \frac{1}{\sigma_N^2}\cdot \textbf{Variance}\,[f(x,y)] \\
&= \frac{\sigma_R^2}{\sigma_N^2}
\end{aligned}$$

Thus, the roughness measure is defined as the ratio of the variance of variance window R to the variance of the normalization window N.

This improved roughness measurement was implemented and tested on all mammogram image pairs with various combinations of variance window sizes from 17 to 31 pixel square and normalization window sizes from 51 to 101 pixel square. Based on visual interpretation of the results, it is found

that a variance window size of 25 pixel square and a normalization window size of 51 pixel square yielded satisfactory results. The response of this measurement on the test Plate 2.1 is depicted in Plate 4.4 with the same convention used as in Plate 4.3. The locally normalized roughness measure produces a much stronger response at the suspicious tumor location than the globally normalized measure.

The locally normalized absolute variance difference measurement is effective at detecting small tumors on a relatively homogeneous background if the corresponding area in the opposite image is smooth. The response peak does not coincide with the center of the tumor as this measurement yields a stronger response to edges than to the center of tumors. Also, this measurement does not detect any difference when corresponding locations have the same variance within the variance window and normalization window such that this ratio measurement is approximately equal to 1, even though these two locations might have completely different overall variance values. Furthermore, it produces no response to large tumors (> 40x40 pixel) with a smooth center such as circumscribed lesions if the opposite breast image has a homogeneous background at the corresponding location.

Observation of all test cases showed that the majority of large tumors were not detected or produced a response only in the edge area. To overcome this deficiency, an additional measurement was developed and is presented in the next section.

4.4 Brightness-to-Roughness Ratio

Typical tumors that are missed by the locally normalized variance measure have a bright and relatively homogeneous tumor center. By comparing the brightness and roughness of this type of tumor to other breast tissue, such as glandular duct and fatty tissue, it was found that they are drastically different. The summary of comparison results are shown in Table 4.1. By defining a measurement as the ratio of brightness and roughness, the tumor is clearly distinguishable from glandular and fatty tissue as illustrated in Table 4.1. Thus, a measurement such as this can overcome the deficiency of the locally normalized variance difference measurement.

	Tumor	Glandular	Fatty
Brightness (μ)	HIGH	MEDIUM - HIGH	LOW
Roughness (σ)	LOW	HIGH	LOW
μ/σ	HIGH	LOW	LOW

Table 4.1: Comparison of Brightness and Roughness on various breast tissue.

The brightness to roughness ratio measurement is once again computed on a windowing system. The brightness measure can be expressed in terms of the mean grey-level value within the window region considered. Since the variance measurement is a first-order statistic measure, it does not reflect the spatial relationship between pixels within the window. For example, the following two subimages have the same variance value.

2	2	0	0
2	2	0	0
2	2	0	0
2	2	0	0

2	0	2	0
0	2	0	2
2	0	2	0
0	2	0	2

Therefore, variance is not used to represent the roughness in this measurement.

Since the roughness measure depends not only on the amplitude of grey-level changes, but also on the spatial frequency of changes, a running sum of absolute grey-level difference between neighboring pixels in both vertical and horizontal directions is used to define the roughness measure of an image window. Mathematically, this roughness measure for a nxn window W, is defined as follows:

$$\sigma = \sum_{j=0}^{255} j^2 \cdot P_d(j)$$

where:

$$P_d(j) = \frac{1}{n^2} \sum_{x,y \in W} [H_j(x,y) + V_j(x,y)]$$

$$H_j(x,y) = \begin{cases} 1 & \text{if } |f(x,y) - f(x+1,y)| = j \\ 0 & \text{otherwise} \end{cases}$$

$$V_j(x,y) = \begin{cases} 1 & \text{if } |f(x,y) - f(x,y+1)| = j \\ 0 & \text{otherwise} \end{cases}$$

The brightness to roughness ratio measure is finally defined as:

$$\mathbf{BR} = \frac{\mu^2}{(1+\sigma)}$$

where μ is the average grey level in the window W. In order to reduce computational cost, images were first reduced into 256x256 pixels. This measurement was then applied to all test cases with various window sizes from 11 to 51 pixel square. A window size of 11 pixel square was selected based on visual judgment of the results. One of the results is depicted in Plate 4.5. The top half of the plate shows the original transformed images and the response of absolute difference of this measurement is shown on the bottom-left of the plate. The red circle indicates the tumor location indicated by the radiologist.

The results show that strong responses were produced for the majority of suspicious tumor areas with a bright and relatively homogeneous center.

4.5 Directionality

Directionality was one of the measures used by Hand and Semmlow *et al.* [HSAA79, SSA+80] in detecting suspicious areas in xeromammograms. Although they gave no motivation for selecting directionality as one of the parameters in their detection algorithm, we believe that this directionality measure is aimed at detecting either spicules of stellate lesions or deformations of blood vessels and glandular ducts caused by the growth of breast

tumors. They measured directionality as scaled sum of intensity changes between neighboring pixels with a fixed spatial distance and direction over a small window [HSAA79, SSA+80]. One example of their directionality measure is given below:

$$\text{Horizontal directionality} = \frac{1}{90}\sum_{i=1}^{9}\sum_{j=1}^{10}|f(i,j) - f(i+10,j)|$$

However, image texture directionality can be better described in the power spectrum as complete information on the grey-level variation for the entire range of spatial frequency and direction is available in the power spectrum.

4.5.1 Fourier Transform

The discrete Fourier transform of an image, $f(x,y)$ is defined as follows:

$$F(u,v) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\exp\left[-2\pi i\left(\frac{mu}{M} + \frac{nv}{N}\right)\right]$$

The transform expresses an image function as a sum of sine waves of different frequency and phase angle. Low spatial frequencies account for the "slowly varying" grey-levels in an image; whereas the high spatial frequencies are associated with "quickly varying" information [BB82, p.25].

Image texture can be described by the power spectrum of an image. The power spectrum represents the distribution of the magnitude of grey-level variation at various frequencies and phase angles and is defined as:

$$P(u,v) = [F(u,v)\cdot F^{*}(u,v)]^{1/2}$$

where $F^*(u,v)$ is the complex conjugate of $F(u,v)$. Various textural features can be extracted from the power spectrum. For instance, coarse textures are indicated by a power spectrum with strong low-frequency components that are associated with large element size; whereas, fine textural patterns have high values at high frequencies. In addition, texture directionality can also be described by the power spectrum. If a texture pattern contains features which are highly oriented in one direction, then the power is concentrated in a single direction in the power spectrum [Lev85, p.444]. The direction is perpendicular to the direction of the oriented features in the image; for example, horizontal streaks in an image will result in strong vertical components in the power spectrum.

In order to simplify the measurement of textural directionality from the two-dimensional power spectrum, it is necessary to break up the power spectrum into wedges centered at the origin [Lev85, p.444]. Each of these wedges tends to measure the angular sensitivity of the range of phase angle it represents. To simplify the measurement of horizontal and vertical direction of mammogram features, the power spectrum of each sub-window of an image is divided into 4 wedges as illustrated in Figure 4.2. The horizontal and vertical directionality is defined as the normalized total power in regions H and V respectively. Mathematically, given the power spectrum $P(i,j)$, directionality

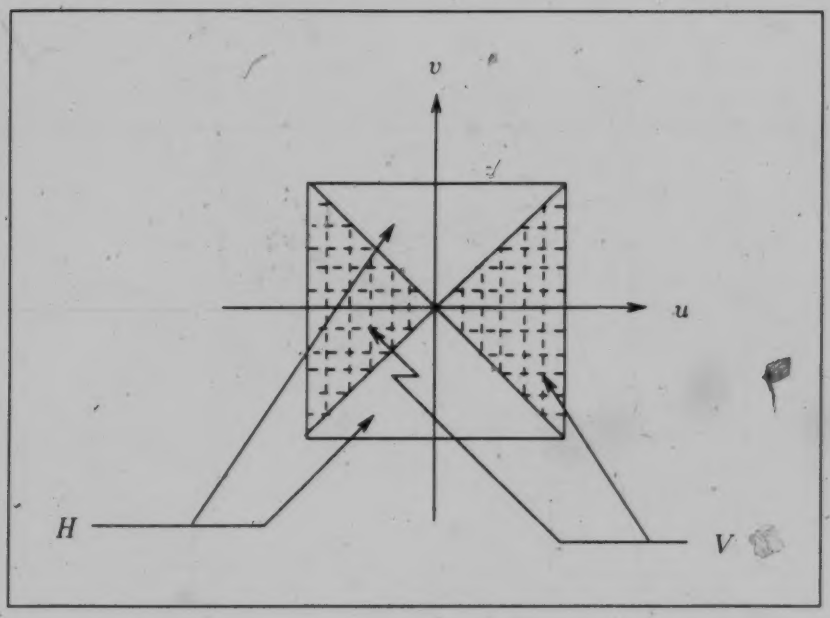


Figure 4.2: Subdivision of power spectrum for directionality.

is defined as:

$$D_H = \frac{\sum_{i,j \in H} P(i,j)}{\sum_{\forall i,j} P(i,j)}$$

$$D_V = \frac{\sum_{i,j \in V} P(i,j)}{\sum_{\forall i,j} P(i,j)}$$

The power spectral method is a computationally expensive method; together with the overlapping window technique, it is not feasible to analyze images in a full scale 512x512 pixels in a mass screening program. Hence, the images were first reduced to a resolution of 256x256 pixels. Moving window sizes of 8, 16, 32 pixel square were tested on all test cases. The two-dimensional Fast Fourier Transform routine was used to compute the Fourier

transform of each window of the breast area. Breast boundary pixels within the width of window size were ignored to eliminate the strong directionality response at the breast boundary. Based on the periodicity properties of Fourier transform where:

$$F(-u,-v) = F(M-u, N-v)$$

the directionality measures can be computed from half of the power spectrum; thus, it saves a considerable amount of computation.

After observing the results from all test cases, it was found that a window size of 16x16 pixels produced the best response for the total absolute difference of the vertical and horizontal directionality measures. However, these results also indicate that the response is dominated by the natural asymmetries of largely oriented features, such as blood vessels and glandular ducts; hence signaling a large number of false suspicious areas. One example of this measurement's response is depicted in Plate 4.6. This type of response behavior is caused by two major factors. First of all, the majority of breast tumors are not as highly oriented as blood vessels or glandular ducts. In addition, oriented patterns of breast tumor, such as spicules of stellate lesion, are barely, if at all, visible in digitized mammograms. Thus, typical breast tumors do not produce a strong response to these directionality measures. Second, anatomic structures such as blood vessels and ducts could not be aligned precisely due to the limitations of geometric transformation used.

Hence, natural asymmetries of these highly oriented features are unavoidable and they lead to responses of the above directionality measures.

For these reasons, the directionality measures were found to be not suitable for detecting mammographic asymmetries caused by breast tumors. However, this measurement can be used as a weighting factor in the combination of all other mammographic asymmetry measurements to reduce the number of false alarms by suppressing the responses due to natural asymmetries. Details of this method are presented in the next section.

4.6 Formation of Asymmetry Measurement

As suggested earlier in this chapter, mammographic asymmetry should be defined based on the difference of a set of textural features as there is no single measurement that can adequately describe mammographic appearance. The more textural features are included in this set, the more likely it is that the mammographic appearance can be described adequately by the set. However, for efficiency reasons, it is not feasible to use a large set of textural features for defining mammographic asymmetry. In addition, a larger number of textural features does not guarantee a more precise asymmetry detection. This is because a large number of false alarms results from the accumulation of unavoidable false alarms for each feature measurement. Therefore, only normalized digital subtraction, locally normalized variance and brightness-

to-roughness ratio were used to detect mammographic asymmetries.

Due to the lack of a sufficient number of test cases, no quantitative studies were done to analyze the performance of each feature measurement in detecting mammographic asymmetry caused by breast tumor growth. Thus, it would be inappropriate to use a weighted linear combination method to combine these three measurements in an overall asymmetry detection measurement as there is no adequate information for defining the weighting factors. Therefore, a simple linear combination method with an equal weighting factor was used to combine the feature measurements.

The final result is formed by combining the responses from the normalized digital subtraction, locally normalized variance difference, and brightness-to-roughness ratio difference in previous sections. In addition, the response to directionality measure on both transformed images was used to form a weighting factor for the above three measurements in order to suppress any natural mammographic asymmetry of oriented pattern, such as blood vessels or glandular ducts. Since all mammograms were digitized with the nipple pointing upward, the majority of the oriented patterns lay within a phase angle region covered by the vertical directionality measure defined in Section 4.5. Thus, only the vertical directionality response from both images was used to define this weighting factor. Further, a power function was used to ensure that only mammographic asymmetries due to highly oriented patterns are removed. Let $D_{V,L}(x,y)$ and $D_{V,R}(x,y)$ be the vertical direc-

tionality on the left and right geometrically transformed mammogram image respectively. The weighting factor for the linear combination of the three feature measurements is defined as:

$$W(x,y) = \begin{cases} 0 & \text{if } P(x,y) > 1 \\ 1 - P(x,y) & \text{otherwise} , \end{cases}$$

where:

$$P(x,y) = \frac{D_{V,L}(x,y)^P + D_{V,R}(x,y)^P}{255^P}$$

Also, let $S(x,y)$, $R(x,y)$ and $BR(x,y)$ represent the responses of normalized digital subtraction, locally normalized variance difference and brightness-to-roughness ratio difference respectively; the linear combination of these measurements' responses with weighting factor can be written as follows:

$$CR(x,y) = [S(x,y) + R(x,y) + BR(x,y)] * W(x,y)$$

Various P values were tested out for $P(x,y)$ in the weighting factor and the value 5 was selected as it produced the best result. An example of this combined response of the test case Plate 2.1 is depicted in Plate 4.7.

4.7 Selection of Suspicious Areas

The response to the combined mammographic asymmetry measurement is in turn a grey-level image as illustrated in the previous section. We now have to devise a method for selecting all the suspicious areas marked by an

expert radiologist from this image. Since the value in the combined response image represents the degree of suspiciousness at its location, it is obvious that one can select the location with peak response value as a suspicious area. However, this naive selection method is not suitable for our application, because multiple suspicious areas might be present in a pair of mammogram images due to multiple tumor growths. Furthermore, prior knowledge of the exact number of suspicious areas is not available; hence, selecting a certain fixed number of suspicious areas is not viable either. A thresholding method seems to be appropriate for this situation.

A single threshold could either produce too many false suspicious areas or miss some of the true suspicious areas. Thus, a two-stage thresholding method is designed for selecting suspicious areas from the combined response image. Before the selection can begin, the combined response image from the previous section is first convolved with a 3x3 low-pass filter to remove any noise due to errors from previous processing operations.

At the first stage of this method, a percentile method is used to determine a threshold for the filtered response image. Based on the assumption that suspicious areas occupy, at most, a fixed percentage of total breast area, a threshold value T_1 is chosen such that q_1 percent of the non-zero response value locations in the filtered response image having a response value higher than T_1 are considered as primary suspicious areas.

Following the first-stage thresholding of the response image, a blob-coloring method [BB82, p.151] is applied to the result image to identify clusters of response locations. Each of these clusters of response locations forms a suspicious area and a bounding circle is determined to represent its location and size. The center coordinate, radius and mean response value of each suspicious area are then stored in a list for further examination.

While the first-stage thresholding process determines the size of each suspicious area, the second-stage is aimed at eliminating false suspicious areas from the previous stage based on the mean response value of each suspicious area. Another threshold, T_2, is determined as q_2 percent of the maximum mean response value of all primary suspicious areas. Thus, a primary suspicious area from the first-stage thresholding which has a mean response value higher than T_2 is considered a true suspicious area; otherwise, it is eliminated from the list as a false suspicious area.

4.7.1 Experimental Results

In determining the value of the parameter q_1 used in the percentile method of the first-stage thresholding, the procedure described was applied to all ten test cases. A value of $q_1 = 5$ produced the best result; and hence, it was selected as the parameter value for the percentile method in the first-stage thresholding method.

Since one cannot expect the mammographic asymmetry detection algorithm to produce suspicious areas perfectly matching those diagnosed by expert radiologists, some criterion for determining a positive detection must be established. The objective of our system is to detect all areas marked as "suspicious" by the radiologist. An area marked as "suspicious" by the system is considered a positive detection if at least 50% of the area overlaps with an area marked as suspicious by the radiologist; otherwise, it is considered a miss.

It is obvious that there is a tradeoff between inclusion of false suspicious areas and exclusion of true suspicious areas when selecting the value for parameter q_2 in the second-stage of the thresholding method. A low q_2 value will produce a large number of false suspicious areas; whereas, a high q_2 value might exclude some of the true suspicious areas. Using three different values for q_2, namely 70, 80 and 90, the second-stage thresholding procedure was applied to the lists of primary suspicious areas generated by the first-stage process on the ten test cases which contain a total of 13 true suspicious areas, as diagnosed by an expert radiologist. The result of comparing the performance of the above three parameter values for q_2 is summarized in Table 4.2. This table shows that the lower the q_2 value used, the higher the number of false suspicious areas detected. However, when a value of 90 was used for q_2, the hit-rate dropped from 92% (q_2 = 80%) to 69%. Based on these results, the value of 80 was selected as the parameter value for q_2 and

the list of size and location of suspicious areas diagnosed by the radiologist, together with those detected by computer, are summarized in Table 4.3. A few examples of the final result of the mammographic asymmetry method proposed in this thesis are depicted in Plate 4.8 and 4.9. The convention of these plates is that the top half is the original geometrically transformed mammogram image pair and the bottom-left is the final response image of the linear combination of the three feature measurements from Section 4.6. The red circles are the true suspicious areas marked by the radiologist and the blue circles are the suspicious areas mapped by the mammographic asymmetry detection program.

q_2 (%)	Hit-Rate (%)	Total False Suspicious Areas Detected	Average Number of False Suspicious Area per case
70	92	99	9.9
80	92	49	4.9
90	69	16	1.6

Table 4.2: Comparison of performance of different q_2 value in second stage thresholding procedure.

CASE NO.	Coordinate of Suspicious Areas						Total False Suspicious Areas
	Radiologist			Computer			
	x	y	radius (pixel)	x	y	radius (pixel)	
19010(ML)	254	159	10	254	146	16	13
19010(CC)	344	196	10	338	192	6	2
	140	216	10	132	212	10	
	160	249	20	MISSED			
140496(CC)	278	323	12	276	328	20	6
	365	364	30	364	364	20	
147336(CC)	190	277	20	186	282	10	3
49746(CC)	286	382	50	298	394	38	0
128686(ML)	166	310	35	172	328	30	3
166468(CC)	306	228	20	306	244	16	4
166468(ML)	296	183	30	296	172	24	5
120610(CC)	336	339	12	334	342	14	9
196348(ML)	229	180	25	222	190	12	4

Table 4.3: Comparison of computer's and radiologist's interpretation on ten test cases.

Plate 4.1: Result of Locally Normalized Digital Subtraction

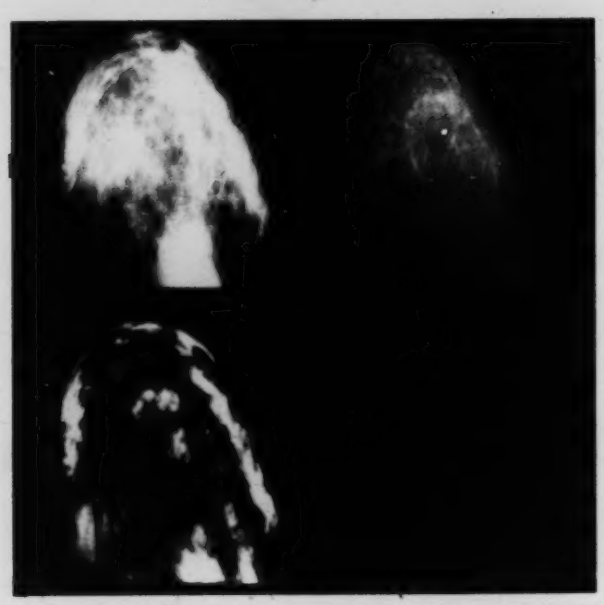


Plate 4.2: Result(case:49746cc) of Locally Normalized Digital Subtraction

Plate 4.3: Result of Globally Normalized Variance Difference

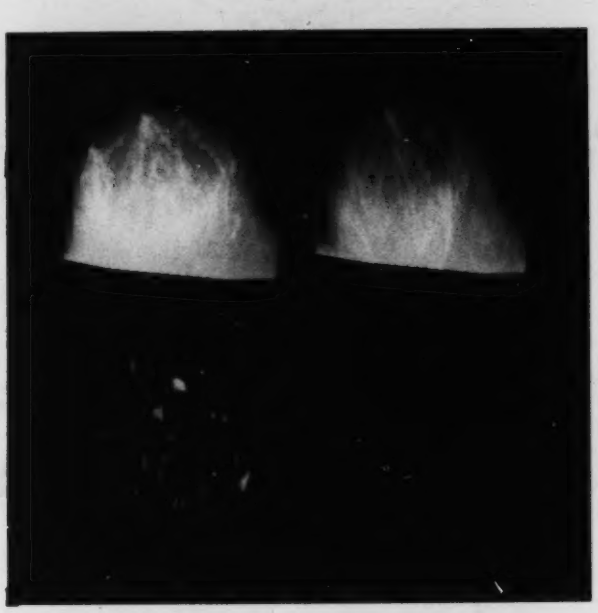


Plate 4.4: Result of Locally Normalized Variance Difference

Plate 4.5: Result of Brightness-to-Roughness Ratio Difference

Plate 4.6: Result of Directionality Difference

Plate 4.7: Result of Combined Responses on test case: Plate 2.1

Plate 4.8: Final Result of Suspicious Areas Detected on Case:19010ML

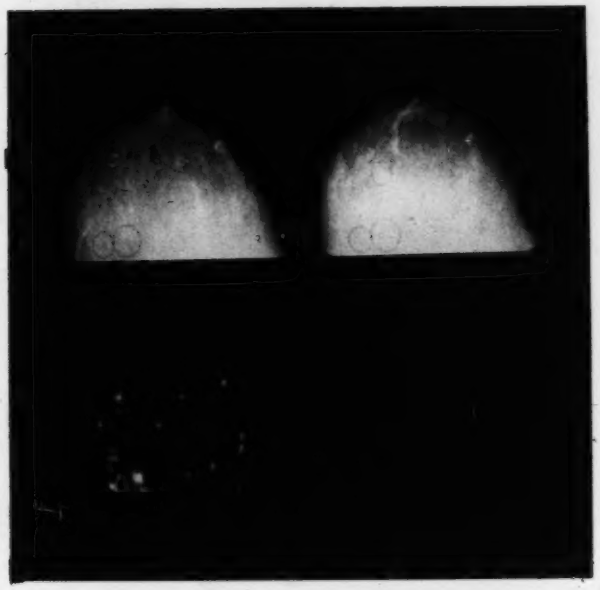


Plate 4.9: Final Result of Suspicious Areas Detected on Case:19010CC

Chapter 5

Conclusions and Future Research

The objective of this thesis was to design and implement an automated breast tumor detection algorithm using the mammographic asymmetry approach. The low image quality of mammograms, the wide variation in mammographic appearance of the human breast, and the existence of natural asymmetries between a pair of identical view mammograms make the problem of detecting mammographic asymmetries caused by growth of breast tumor a difficult one.

The research first conducted a background study on the medical aspect of breast tumors, followed by a survey of diagnostic imaging techniques commonly used in clinical practice. Mammographic diagnosis procedure was then discussed in depth.

Based on the interpretation procedure used by expert radiologists, we discovered that there are basically two major approaches used by different researchers in designing their automated breast tumor detection methods from mammograms, namely the *Individual Tumor Type Approach* and *Mammographic Asymmetry Approach*. A review of recent research projects in automation of mammographic detection of breast tumor using the above two approaches was presented. The advantages of using the mammographic asymmetry approach in designing automated mammographic breast tumor detection method were studied and discussed. All processes involved in the mammographic asymmetry detection method were then identified and discussed in detail.

The first step of the mammographic asymmetry detection method was aimed at breast area extraction. A global thresholding segmentation method with manual selection of threshold value was adopted. An unweighted averaging filter of size 11x11 pixels was used to smooth the fuzzy boundary of the binary images produced by global thresholding. This method yielded satisfactory results on all test cases.

One of the major problems involved in developing an automated mammographic asymmetry detection method was to devise a simple but adequate geometric transformation method that can align a pair of identical view mammograms and eliminate natural discrepancies in size and shape. The initial requirement for any geometric transformation operation is to obtain some

visible and unambiguously defined corresponding features from both images. However, it is difficult, if not impossible, to reliably detect such features without human assistance. Thus, three points along the breast boundary, namely the nipple and intersection points between breast boundary and back boundary, were selected as the control points for the geometric transformation process. Based on the characteristic curvature property of these control points, the cubic B-spline corner detection method proposed in [MY87] was adopted and modified to detect these points in each mammogram. In addition, an interactive program allowing the user to enter the nipple position was written for the case where the nipple profile does not appear in the digitized image.

In the course of pursuing a geometric transformation for aligning mammogram images, it was found that existing methods were either too simple and did not yield adequate results [HSAA79, SSA+80] or were too complicated to be used effectively and feasibly in a mass screening program [ZG88]. A local geometric transformation method for transforming the breast area in three sub-regions was proposed and implemented. This proposed method produced adequate results for all test cases.

Another major topic of this thesis was to find a set of a minimum number of measurements to detect mammographic asymmetries caused by breast tumor growth. A number of different measurements were studied and implemented, and four measurements were selected to form the final mammographic asymmetry measurement. These measurements include locally nor-

malized digital subtraction, locally normalized variance difference, brightness-to-roughness ratio difference and directionality measure. A vertical directionality measure was used as a weighting factor to suppress responses due to natural asymmetries of highly oriented structures, such as blood vessels and glandular ducts.

The response of the combined asymmetry measurements was then thresholded using a percentile method and followed by a blob-coloring method to find clusters of responses which form the primary suspicious areas. After the size, center of bounding circle and mean response value of each primary suspicious area were determined, a second thresholding procedure was used to eliminate false suspicious areas from the primary suspicious areas based on the mean response value within each suspicious area.

The entire mammographic asymmetry detection method was implemented and applied to ten test cases. The results showed that 92% of all suspicious areas were correctly identified. Furthermore, a significantly lower false alarm rate (4.9/case) compared to Hand *et al.*'s studies [HSAA79] (53.6/case) was achieved.

The research of this thesis provides a complete overall framework in designing an automated breast tumor detection method from conventional film mammograms using mammographic asymmetry approach. It also provides an initial step towards the development of an automated mammogram screening method which can be used in a mass screening program.

5.1 Future Directions

Given that this is an initial attempt to develop automated breast tumor detection using the mammographic asymmetry approach, the method presented in this thesis produced quite encouraging results. However, the method is still far from being reliable enough to be used in a mass screening program. Further research is required to improve this method.

First of all, the stability of this method must be tested on a large number of sample cases. This will not only provide more insight into selecting more appropriate parameter values in the suspicious area selection process; but will also provide more complete knowledge for determining weighting factors for each of the three asymmetry measurements in the linear combination process.

Even though the proposed geometric transformation method produced adequate results in aligning identical view mammogram images; it does not eliminate natural asymmetries within the breast area. Future work should concentrate on detecting corresponding anatomic structures, such as blood vessels and glandular ducts, and incorporate these structures as control points in the geometric transformation process. This will further eliminate natural asymmetries between a pair of mammograms and provide a more accurate alignment.

Finally, this research has only investigated a small number of measurements for mammographic asymmetry measurement. The study to find the best mammographic asymmetry measurement is far from complete and exhaustive. Future research should concentrate on the study of other measurements and their applicability in measuring mammographic asymmetry.

Bibliography

[AG72] Laurens V. Ackerman and Earl E. Gose. Breast lesion classification by computer and xeroradiography. *Cancer*, 30:1025–1035, 1972.

[BB82] Dana H. Ballard and Cristopher M. Brown. *Computer Vision.* Prentice-Hall, Inc., Englewood Cliffs, New Jersey, 1982.

[CK72] C.K. Chow and T. Kaneko. Automatic boundary detection of left ventricle from cineangiograms. *Computers and Biomedical Research*, 5:388–410, 1972.

[Col87] Catherine Cole-Beuglet. Ultrasound. In Lawrence W. Bassett and Richard H. Gold, editors, *Breast Cancer Detection: Mammography and Other Methods in Breast Imaging*, pages 153–168. Grune and Stratton, Inc., Orlando, Florida, 1987.

[CSF+80] C.H. Joseph Chang, Justo L. Sibala, Steven L. Fritz, Samuel J. Dwyer III, Arch. W. Templeton, Fritz Lin, and William R. Jewell. Computed tomography in detection and diagnosis of breast cancer. *Cancer*, 46:939–946, August 1980.

[CTH84] R.W. Conners, M.M. Trivedi, and C.A. Harlow. Segmentation of a high-resolution urban scene using texture operators. *Comput. Vision Graphics and Image Process.*, 25:273–310, 1984.

[DMH76] S.J. Dwyer III, R.W. McLaren, and C.A. Harlow. Computer-aided diagnosis of breast cancer from thermography. In C.H.

Chen, editor, *Pattern Recognition and Artificial Intelligence*, pages 233–247. Academic Press, Inc., New York, USA, 1976.

[DS83] Carl J. D'Orsi and Edward H. Smith. Breast imaging. In Carl J. D'Orsi and Richard E. Wilson, editors, *Carcinoma of the Breast: Diagnosis and Treatment*, chapter 6, pages 95–162. Little, Brown and Company, Boston, 1983.

[Fei87] Stephen A. Feig. Xeromammography. In Lawrence W. Bassett, editor, *Breast Cancer Detection: Mammography and Other Methods in Breast Imaging*, pages 89–109. Grune and Stratton Inc., Orlando USA, 1987.

[Fis85] Edwin R. Fisher. What is early breast cancer? In J. Zander and J. Baltzer, editors, *Early Breast Cancer*, pages 1–13. Springer-Verlag, Berlin, Germany, 1985.

[GB83] Richard H. Gold and Lawrence W. Bassett. Mammography: History and state of the art. In Stephen A. Feig and Robert McLelland, editors, *Breast Carcinoma: Current Diagnosis and Treatment*, pages 95–98. Masson Publishing USA, Inc., New York, 1983.

[GKR77] John J. Gisvold, Philip R. Karsell, and David F. Reese. Computerized tomographic mammography. In Wende Westinghouse Logan, editor, *Breast Carcinoma: The Radiologist's Expanded Role*, pages 219–238. John Wiley and Sons, Inc., New York, 1977.

[Gos83] Ardeshir Goshtasby. *A Symbolically-Assisted Approach To Digital Image Registration With Application in Computer Vision.* PhD thesis, Michigan State University, Michigan, USA, 1983.

[Har79] R.M. Haralick. Statistical and structural approaches to texture. In *Proceedings of the IEEE*, pages 786–804. IEEE, May 1979.

[HSAA79] William Hand, John L. Semmlow, Lauren V. Ackerman, and Frank S. Alcorn. Computer screening of xeromammograms: A technique for defining suspicious areas of the breast. *Computers and Biomedical Research*, 12:445–460, 1979.

[HSD73] R.M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on System, Man, and Cybernetics*, SMC-3(6):610–621, November 1973.

[HWG87] Dong-Chen He, Li Wang, and Jean Guibert. Texture feature extraction. *Pattern Recognition Letters*, 6:269–273, 1987.

[KOS77] Carolyn Kimme, Bernard J. O'Loughlin, and Jack Sklansky. Automatic detection of suspicious abnormalities in breast radiographs. In A. Klinger, K.S. Fu, and T.L. Kunii, editors, *Data Structures, Computer Graphics, and Pattern Recognition*, pages 427–447. Academic Press, Inc., New York, 1977.

[Lai88] Shuk-Mei Lai. A technique for automated detection of breast tumors in mammograms. Master's thesis, University of Alberta, Edmonton, Canada, 1988.

[Les84] Richard G. Lester. The contributions of radiology to the diagnosis, management, and cure of breast cancer. *Radiology*, 151(1):1–7, 1984.

[Lev85] Martin D. Levine. *Vision In Man And Machine*. McGraw-Hill Book Co., New York, 1985.

[LJ87] Wende W. Logan and Joyce A. Janus. Screen-film mammography. In Lawrence W. Bassett, editor, *Breast Cancer Detection: Mammography and Other Methods in Breast Imaging*, pages 75–88. Grune and Stratton Inc., Orlando, USA, 1987.

[LR80] Marc S. Lapayowker and George Revesz. Thermography and ultrasound in detection and diagnosis of breast cancer. *Cancer*, 46:933–938, August 1980.

[Mar83] John E. Martin. Breast imaging techniques: Mammography, ultrasonography, computed tomography, thermography, and transillumination. *Radiologic Clinics of North America*, 21(1):149–153, March 1983.

[McL84] Robert McLelland. Mammography 1984 - challenge to radiology. *AJR*, 143:1–4, July 1984.

[MGE+84] Tom R. Miller, Kenneth J. Goldman, David M. Epstein, Daniel R. Biello, Kondapuram S. Sampathkumaran, Bharath Kumar, and Barry A. Siegel. Improved interpretation of gated cardiac images by use of digital filters. *Radiology*, 152:795–800, 1984.

[MY87] Gerard Medioni and Yoshio Yasumoto. Corner detection and curve representation using cubic b-splines. *Computer Vision, Graphics, and Image Processing*, 39:267–278, 1987.

[RK82] Azriel Rosenfeld and Avinash C. Kak. *Digital Picture Processing*, volume 1 and 2. Academic Press, Inc., Orlando, Florida, second edition, 1982.

[SK83] Norman Sadowsky and Daniel B. Kopans. Breast cancer. *Radiology Clinics of North America*, 21(1):51–65, March 1983.

[SSA+80] John L. Semmlow, Annapoorin Shadagopappan, Laurens V. Ackerman, William Hand, and Frank S. Alcorn. A fully automated system for screening xeromammograms. *Computers and Biomedical Research*, 13:350–362, 1980.

[ST82] Elizabeth L. Schmitt and Barbara Threatt. Tumor location and detectability in mammographic screening. *AJR*, 139:761–765, October 1982.

[TD85] Laszlo Tabar and Peter B. Dean. Basic principles of mammographic diagnosis. *Diagn. Imag. clin. Med.*, 54:146–157, 1985.

[TD87] Laszlo Tabar and Peter B. Dean. The control of breast cancer through mammography screening - what is the evidence? *Radiologic Clinics of North America*, 25(5):993–1005, September 1987.

[Urb85] Jerome A. Urbab. Bilateral breast cancer revealed by biopsy of the opposite breast. In J. Zander and J. Baltzer, editors, *Early Breast Cancer*, pages 33–38. Springer-Verlag, Berlin, Germany, 1985.

[WEM+67] Fred Winsberg, Milton Elkin, Josiah Macy, Victoria Bordaz, and William Weymouth. Detection of radiographic abnormalities in mammograms by means of optical scanning and computer analysis. *Radiology*, 89:211–215, August 1967.

[ZG88] Xiaohua Zhou and Richard Gordan. Geometric unwarping for digital subtraction mammography. In *Proceeding of Vision Interface '88*, pages 25–30, Edmonton, Canada, June 1988. Canadian Image Processing and Pattern Recognition Society.

Glossary

Areola The pigmented ring of tissue that surrounds the nipple.

Benign Good prognosis; favorable; propitious; not malignant.

Calcification Calcium salts laid down in tissue; usually complex form of calcium phosphates and carbonates; found in both benign and malignant breast diseases.

Cancer A mass of tissue which is malignant, invasive and tending to recur after excision and to metastasize to other tissue.

Cranio-caudal Projection image of the breast obtained by directing x-ray beam through the breast in head to foot direction.

Medio-lateral Projection image of the breast obtained by directing x-ray beam through the breast from medial side to lateral side.

Palpate To examine by touch; to feel, touch.